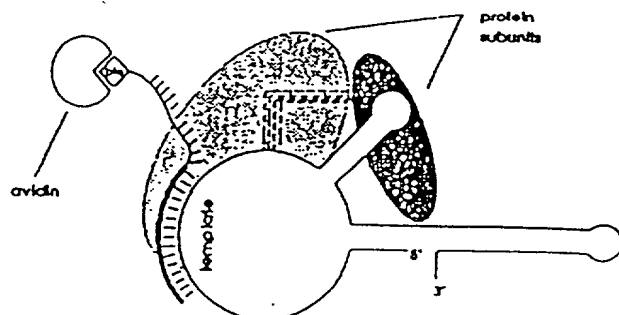


FIGURE 1

PANEL A



PANEL B

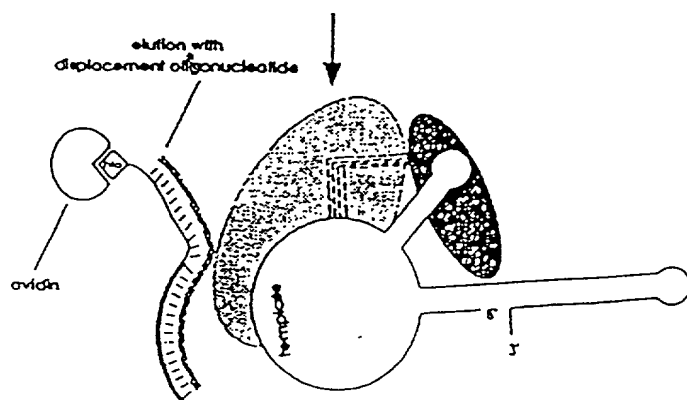


FIGURE 2

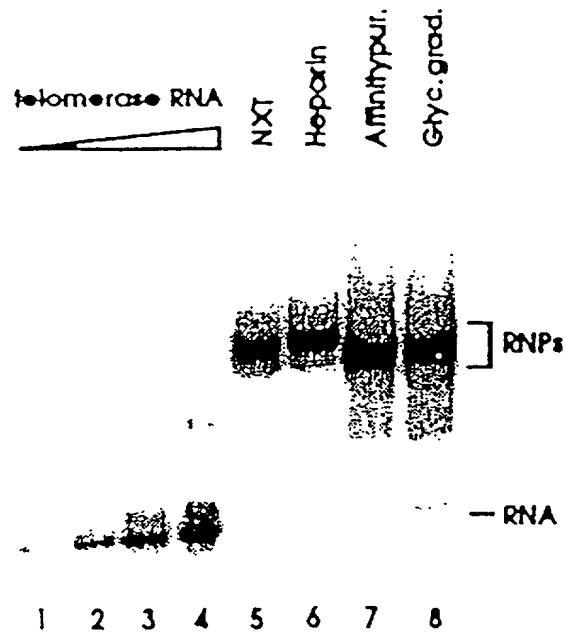


FIGURE 3

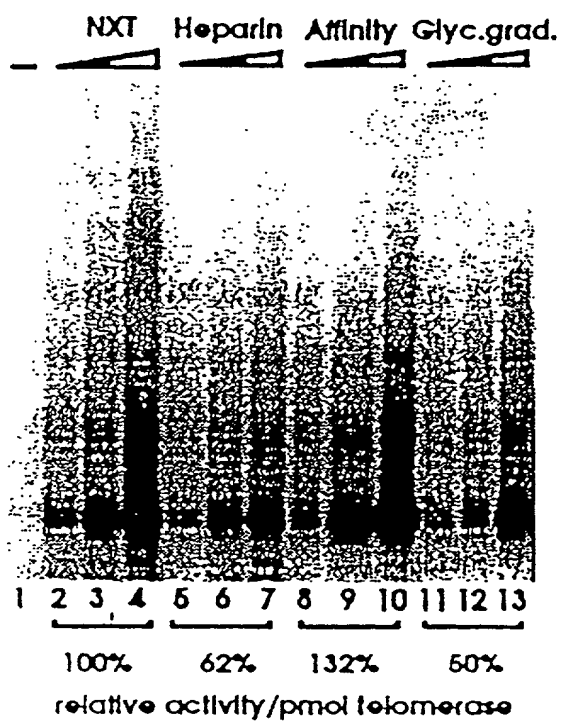


FIGURE 4

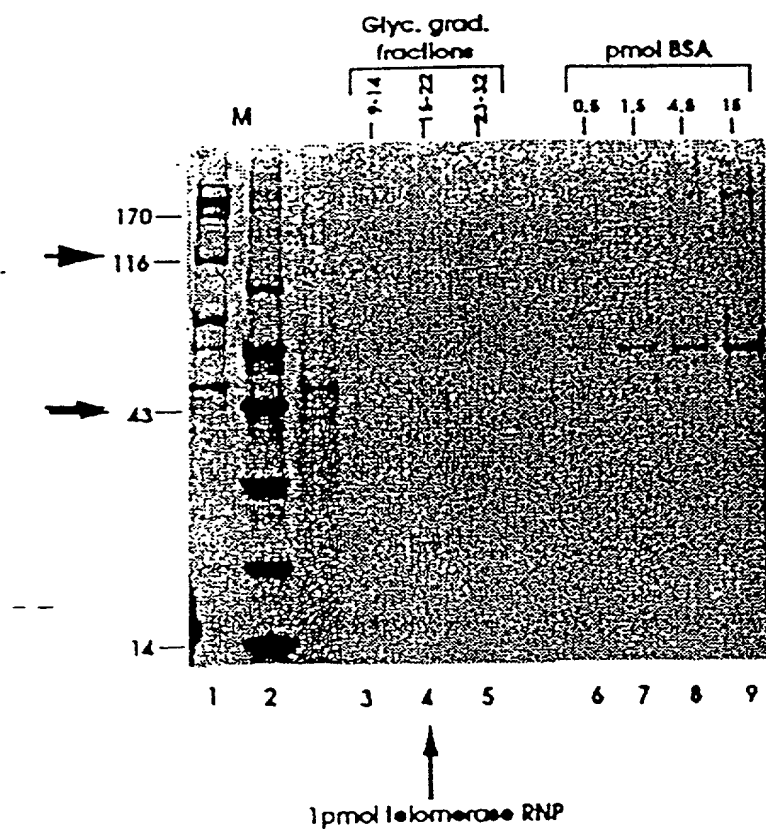


FIGURE 5

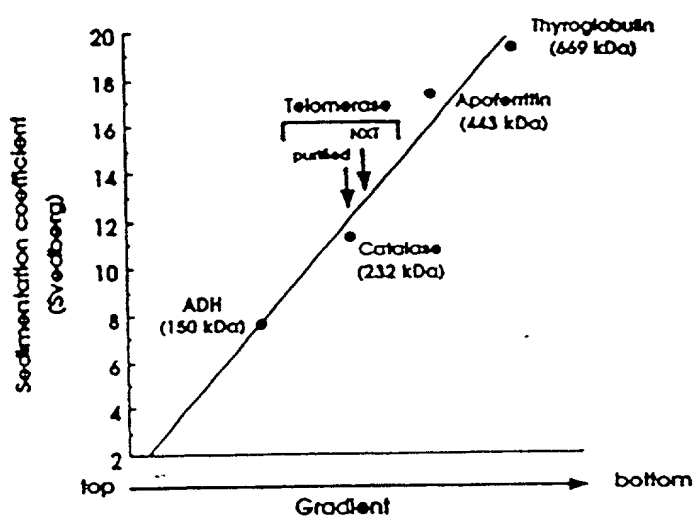


FIGURE 6

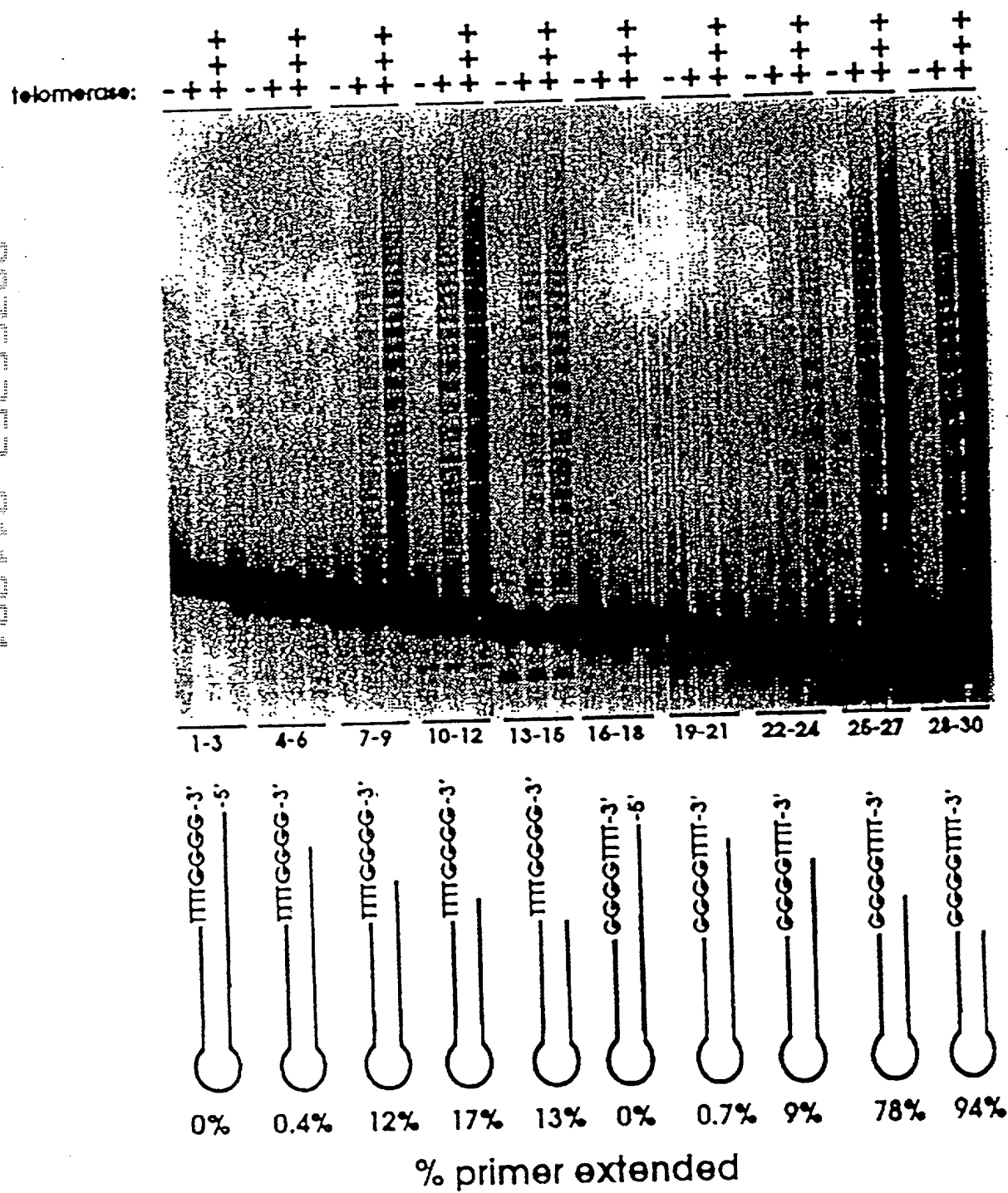


FIGURE 7

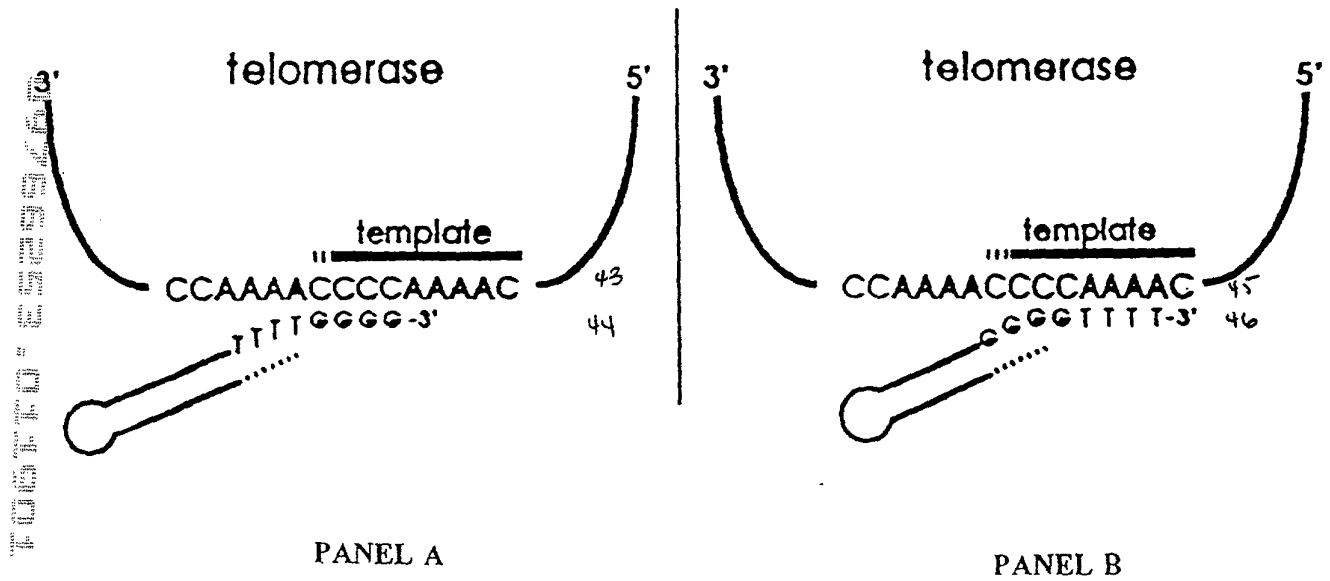


FIGURE 8

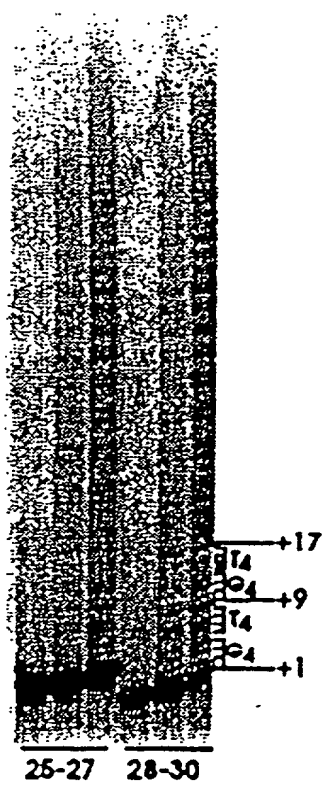


FIGURE 9

1 AAAACCCCAA AACCCCAAAA CCCCTTTTAG AGCCCTGCAG TTGGAAATAT
 51 AACCTCAGTA TTAATAAGCT CAGATTTTAA ATATTAATTA CAAAACCTAA
 101 ATGGAGGTTG ATGTTGATAA TCAAGCTGAT AATCATGGCA TTCCTCAGC
 151 TCTTAAGACT TGTGAAGAAA TTAAAGAAGC TAAAACGTTG TACTCTTGGA
 201 TCCAGAAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA TTATAAAGAT
 251 TTAGAAGATA TTAAAATATT TGCGCAGACA AATATTGTTG CTACTCCACG
 301 AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAA GAAGTATTTT
 351 CAACTGGACT AATGATCGAA CTTATTGACA AATGCTTAGT TGAACCTCTT
 401 TCATCAAGCG ATGTTTCAGA TAGACAAAAA CTTCAATGAT TTGGATTTC
 451 ACTTAAGGGA AATCAATTAG CAAAGACCCA TTTATTAACA GCTCTTTCAA
 501 CTCAAAGCA GTATTTCTTT CAAGACGAAT GGAACCAAGT TAGAGCAATG
 551 ATTGGAAATG AGCTCTCCG ACATCTCTAC ACTAAATATT TAATATTCCA
 601 GCGAATTCT GAAGGAACTC TTGTTCAATT TTGCGGGAAT AACGTTTTTG
 651 ATCATTTGAA AGTCAACGAT AAGTTTGACA AAAAGCAAAA AGGTGGAGCA
 701 GCAGACATGA ATGAACCTCG ATGTTGATCA ACCTGCAAAT ACAATGTCAA
 751 GAATGAGAAA GATCACTTTC TCAACAACAT CAACGTGCCG AATTGGAATA
 801 ATATGAAATC AAGAACCAGA ATATTTTATT GCACTCATTT TAATAGAAAT
 851 AACCAATTCT TCAAAAAGCA TGAGTTTGTG AGTAACAAAA ACAATATTTT
 901 AGCGATGGAC AGAGCTCAGA CGATATTCAC GAATATATTC AGATTTAATA
 951 GAATTAGAAA GAAGCTAAAA GATAAGGTTA TCGAAAAAAT TGCCTACATG
 1001 CTTGAGAAAG TCAAAGATTT TAACTTCAAC TACTATTTAA CAAAATCTTG
 1051 TCCTCTTCCA GAAAATTGGC GGGAACGGAA ACAAAAAATC GAAAATTGA
 1101 TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGA GCTGTTTACG
 1151 TACACAACCTG ATAATAAATG CGTCACACAA TTTATTAATG AATTTTCTA
 1201 CAATATACTC CCCAAAGACT TTTTGACTGG AAGAAACCGT AAGAATTTTC
 1251 AAAAGAAAGT TAAGAAATAT GTGGAACATA ACAAGCATGA ACTCATTAC
 1301 AAAAATTAT TGCTTGAGAA GATCAATACA AGAGAAATAT CATGGATGCA
 1351 GGTTGAGACC TCTGCAAAGC ATTTTATTA TTTTGATCAC GAAAACATCT
 1401 ACGTCTTATG GAAATTGCTC CGATGGATAT TCGAGGATCT CGTCGTCTCG
 1451 CTGATTAGAT GATTTTCTA TGTCACCGAG CAACAGAAAA GTTACTCCAA
 1501 AACCTATTAC TACAGAAAGA ATATTTGGGA CGTCATTATG AAAATGTCAA
 1551 TCGCAGACTT AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT
 1601 GAAGAATGGA AAAAGTCGCT TGGATTTGCA CCTGGAAAAC TCAGACTAAT
 1651 ACCGAAGAAA ACTACTTTCC GTCCAATTAT GACTTTCAAT AAGAAGATTG
 1701 TAAATTCAGA CCGGAAGACT ACAAATTA CTACAAATAC GAAGTTATTG
 1751 AACTCTCACT TAATGCTTAA GACATTGAAG AATAGAATGT TTAAAGATCC
 1801 TTTTGGATTG GCTGTTTTTA ACTATGATGA TGTAATGAAA AAGTATGAGG
 1851 AGTTTGTGTTG CAAATGGAAG CAAGTTGGAC AACCAAACT CTTCTTTGCA
 1901 ACTATGGATA TCGAAAAGTG ATATGATAGT GTAAACAGAG AAAAATATC
 1951 AACATTCTTA AAACTACTA AATTACTTTC TTCAGATTTC TGGATTATGA
 2001 CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA TTCGAAAAAC
 2051 TTTAGAAAGA AAGAAATGAA AGATTATTTT AGACAGAAAT TCCAGAAGAT
 2101 TGCATTGAA GGAGGACAAT ATCCAACCTT ATTCAGTGTT CTTGAAAATG
 2151 AACAAAATGA CTTAAATGCA AAGAAAACAT TAATTGTTGA AGCAAAGCAA
 2201 AGAAATTATT TTAAGAAAGA TAACTTACTT CAACCAGTCA TTAATATTTG
 2251 CCAATATAAT TACATTAAT TTAATGGGAA GTTTTATAAA CAAACAAAAG
 2301 GAATTCCTCA AGGTCTTTGA GTTTCATCAA TTTTGTATC ATTTTATTAT

FIGURE 9 (cont.)

2351 GCAACATTAG AGGAAAGCTC CTTAGGATTC CTTAGAGATG AATCAATGAA
2401 CCCTGAAAAT CCAAATGTTA ATCTTCTAAT GAGACTTACA GATGACTATC
2451 TTTTGATTAC AACTCAAGAG AATAATGCAG TATTGTTTAT TGAGAAACTT
2501 ATAAACGTAA GTCGTGAAAA TGGATTTAAA TTCAATATGA AGAAACTACA
2551 GACTAGTTTT CCATTAAGTC CAAGCAAATT TGCAAAATAC GGAATGGATA
2601 GTGTTGAGGA GCAAAATATT GTTCAAGATT ACTGCGATTG GATTGGCATC
2651 TCAATTGATA TGAAAACCTCT TGCTTTAATG CCAAATATTA ACTTGAGAAT
2701 AGAAGGAATT CTGTGTACAC TCAATCTAAA CATGCAAACA AAGAAAGCAT
2751 CAATGTGGCT CAAGAAGAAA CTAAAGTCGT TTTTAATGAA TAACATTACC
2801 CATTATTTTA GAAAGACGAT TACAACCGAA GACTTTGCGA ATAAAACCTCT
2851 CAACAAGTTA TTTATATCAG GCGGTTACAA ATACATGCAA TGAGCCAAAG
2901 AATACAAGGA CCACTTTAAG AAGAACTTAG CTATGAGCAG TATGATCGAC
2951 TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA
3001 CCTTGTGTGC AATATTAAGG ATACAATTTT TGGAGAGGAG CATTATCCAG
3051 ACTTTTTCTT TAGCACACTG AAGCACTTTA TTGAAATATT CAGCACAAAA
3101 AAGTACATTT TCAACAGAGT TTGCATGATC CTCAAGGCAA AAGAAGCAAA
3151 GCTAAAAAGT GACCAATGTC AATCTCTAAT TCAATATGAT GCATAGTCGA
3201 CTATTCTAAC TTATTTTGA AAGTTAATTT TCAATTTTTG TCTTATATAC
3251 TGGGGTTTTG GGGTTTTGGG GTTTTGGGG

Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

1 MEVDVDNQAD NHGIHSALKT CEEIKEAKTL YSWIQKVICR RNQSQSHYKD
51 LEDIKIFAQT NIVATPRDYN EEDFKVIARK EVFSTGLMIE LIDKCLVELL
101 SSSDVSDRQK LQCFGFQLKG NQLAKTHLLT ALSTQKQYFF QDEWNQVRAM
151 IGNELFRHLY TKYLIFQRTS EGTLVQFCGN NVFDHLKVND KFDKKQKGGA
201 ADMNEPRCCS TCKYNVKNK DHFLNNINVP NWNMKSRTTR IFYCTHFNRN
251 NQFFKKHEFV SNKNNISAMD RAQTIFTNIF RFNRIRKKLK DKVIEKIAYM
301 LEKVKDFNFN YYLTKSCPLP ENWRERKQKI ENLINKTREE KSKYEEELFS
351 YTTDNKCVTQ FINEFFYNIL PKDFLTGRNR KNFQKKVKKY VELNKHელი
401 KNLLLEKINT REISWMQVET SAKHFYYFDH ENIYVLWKKL RWIFEDLVVS
451 LIRCFYVTE QQKSYSKTY YRKNIWVIM KMSIADLKE TLAEVQEKEV
501 EEWKKS LGFA PGKLR LIPKK TTRPIMTFN KKIVNSDRKT TKLTTNTKLL
551 NSHMLKTLK NRMFKDPFGF AVFNYDDVMK KYEEFVCKWK QVGQPKLFFA
601 TMDIEKCYDS VNREKLSTFL KTTKLLSSDF WIMTAQILKR KNNIVIDSKN
651 FRKKEMKDYF RQKFQKIALE GGQYPTLFSV LENEQNDLNA KKTLLIVEAKQ
701 RNYFKKDNLL QPVINICQYN YINFNGKFYK QTKGIPQGLC VSSILSSFYY
751 ATLEESSLGF LRDESMNPEN PNVNLLMRLT DDYLLITTQE NNAVLFIKEL
801 INVSRENGFK FNMKKLQTSF PLSPSKFAKY GMDSVEEQNI VQDYCDWIGI
851 SIDMKTLALM PNINLRIEGL LCTLNLMQNT KKASMWLKKK LKSFLMNNIT
901 HYFRKTITTE DFANKTLNKL FISGGYKYM CAKEYKDHFK KNLAMSSMID
951 LEVSKIIYSV TRAFFKYLCV NIKDTIFGEE HYPDFFLSTL KHFIIEFSTK
1001 KYIFNRVCM I LKAKEAKLKS DQCQSLIQYD A

FIGURE 11

1 CCCCAAACC CCAAACCCC AAAACCCCTA TAAAAAAGA AAAAATTGAG
 51 GTAGTTTAGA AATAAAATAT TATCCCGCA CAAATGGAGA TGGATATTGA
 101 TTTGGATGAT ATAGAAAATT TACTTCCTAA TACATTCAAC AAGTATAGCA
 151 GCTCTGTAG TGACAAGAAA GGATGCAAAA CATTGAAATC TGGCTCGAAA
 201 TCGCCTTCAT TGACTATTCC AAAGTTGCAA AAACAATTAG AGTTCTACTT
 251 CTCGGATGCA AATCTTTATA ACGATTCTTT CTTGAGAAAA TTAGTTTAA
 301 AAAGCGGAGA GCAAAGAGTA GAAATTGAAA CATTACTAAT GTTTAAATAA
 351 AATCAGGTAA TGAGGATTAT TCTATTTTTT AGATCACTTC TTAAGGAGCA
 401 TTATGGAGAA AATTACTTAA TACTAAAAGG TAAACAGTTT GGATTATTC
 451 CCTAGCCAAC AATGATGAGT ATATTAAATT CATATGAGAA TGAGTCAAAG
 501 GATCTCGATA CATCAGACTT ACCAAAGACA AACTCGCTAT AAAACGCAAG
 551 AAAAAGTTTG ATAATCGAAC AGCAGAAGAA CTTATTGCAT TACTATTTCG
 601 TATGGGTTTT ATTACAATTG TTTAGGTAT CGACGGTGAA CTCCCGAGTC
 651 TTGAGACAAT TGAAAAAGCT GTTACAACCT GAAGGAATCG CAGTTCTGAA
 701 AGTTCTGATG TGTATGCCAT TATTTTGTGA ATTAATCTCA AATATCTTAT
 751 CTCAATTTAA TGGATAGCTA TAGAAACAAA CCAAATAAAC CATGCAAGTT
 801 TAATGGAATA TACGTAAAT CCTTTGGGAC AAATGCACAC TGAATTTATA
 851 TTGGATTCTT AAAGCATAGA TACACAGAAT GCTTTAGAGA CTGATTAGC
 901 TTACAACAGA TTACCTGTTT TGATTACTCT TGCTCATCTC TTATATCTTT
 951 AAAAGAAGCA GCGGAAATGA AAAGAAGACT AAAGAAAGAG ATTTCAAAT
 1001 TTGTTGATTC TTCTGTAACC GGAATTAACA ACAAGAATAT TAGCAACGAA
 1051 AAAGAAGAAG AGCTATCACA ATCCTGATTC TTAAAGATTT CAAAAATTCC
 1101 AGGTAAGAGA GATACATTCA TTTAAATTCA TATATTATAG TTTTTCATTT
 1151 CACAGCTGTT ATTTTCTTTT ATCTTAACAA TATTTTTTGA TTAGCTGGAA
 1201 GTAAAAAGTA TCAAATAAGA GAAGCGCTAG ACTGAGGTAA CTTAGCTTAT
 1251 TCACATTCAT AGATCGACCT TCATATATCC AATACGATGA TAAGGAAACA
 1301 GCAGTCATCC GTTTTAAAAA TAGTGCTATG AGGACTAAAT TTTTAGAGTC
 1351 AAGAAATGGA GCCGAAATCT TAATCAAAAA GAATTGCGTC GATATTGCAA
 1401 AAGAATCGAA CTCTAAATCT TTCGTTAATA AGTATTACCA ATCTTGATTG
 1451 ATTGAAGAGA TTGACGAGGC AACTGCACAG AAGATCATT AAGAAATAAA
 1501 GTAACTTTTA TTAATTAGAG AATAAACTAA ATTACTAATA TAGAGATCAG
 1551 CGATCTTCAA TTGACGAAAT AAAAGCTGAA CTAAAGTTAG ACAATAAAAA
 1601 ATACAAACCT TGGTCAAAAT ATTGAGGAAG GAAAAGAAGA CCAGTTAGCA
 1651 AAAGAAAAAA TAAGGCAATA AATAAAATGA GTACAGAAGT GAAGAAATAA
 1701 AAGATTATTT TTTTCAATA ATTTATTGAA AAGAGGGGTT TTGGGGTTTT
 1751 GGGGTTTTGG GG

CCCAAACC CCAAACCCC

[illegible][illegible]

11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100										
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

a
b
c

FIGURE 12 (cont.)

1081 TTAAGAGATTTCAAAAATCCAGGTAAGAGAGATACATTCATTAAAAATTCATATATTATAG 1140
 AATTTCTAAAGTTTTTAAGGTCCATTCTCTCTATGTAAGTAATTTTAAGTATATAATATC

a L K I S K I P G K R D T F I K I H I L * -
 b * R F Q K F Q V R E I H S L K F I Y Y S -
 c K D F K N S R * E R Y I H * N S Y I I V -

1141 TTTTTCATTTCACAGCTGTTATTTTCTTTTATCTTAACAATATTTTGTAGCTGGAA 1200
 AAAAAGTAAAGTCTGCACAATAAAAGAAAATAGAATTGTTATAAAAAACTAATCGACCTT

a F F I S Q L L F S F I L T I F F D * L E -
 b F S F H S C Y F L L S * Q Y F L I S W K -
 c F H F T A V I F F Y L N N I F C L A G S -

1201 GTAAAAAGTATCAATAAGAGAAGCGCTAGACTGAGGTAACCTAGCTTATTCACATTCAT 1260
 CATTTTTCATAGTTTATCTCTCTCGCATCTGACTCCATTGAATCGAATAAGTGTAAAGTA

a V K S I K * E K R * T E V T * L I H I H -
 b * K V S N K R S A R L R * L S L F T F I -
 c K K Y Q I R E A L D C G N L A Y S H S *

1261 AGATCGACCTTCATATATCCAATACGATGATAAGGAAACAGCAGTCATCCGTTTTAAAAA 1320
 TCTAGCTGGAAGTATATAGGTTATGCTACTATTCCTTTGTCGTGCTAGGCAAAATTTT

a R S T F I Y P I R C * G N S S H P F * K -
 b D R P S Y I Q Y D D K E T A V I R F K N -
 c I D L H I S N T M I R K Q Q S S V L K I -

1321 TAGTGCTATGAGGACTAAATTTTAGAGTCAAGAAATGGAGCCGAAATCTTAATCAAAAA 1380
 ATCACCATACTCTGATTAAAAATCTCAGTTCTTTACCTCGGCTTAGAATTAGTTTTT

a * C Y E D * I F R V K K W S R N L N Q K -
 b S A M R T K F L E S R N G A E I L I K K -
 c V L C G L N F * S Q E M E P K S * S K R -

1381 GAATTGCGTCGATATTGCAAAAGAATCGAACTCTAAATCTTTCGTTAATAAGTATTACCA 1440
 CTTAACGCAGCTATAACGTTTTCTTAGCTTGAGATTAGAAAGCAATTATTCATAATGGT

a E L R R Y C K R I E L * I F R * * V L P -
 b N C V D I A K E S N S K S F V N K Y Y Q -
 c I A S I L Q K N R T L N L S L I S I T N -

1441 ATCTTGATTGATTGAAGAGATTGACGAGGCAACTGCACAGAAGATCATTAAGAAATAAA 1500
 TAGAACTAACTAATCTCTAACTGCTCCGTTGACGTGCTCTCTAGTAATTCTTTATTT

a I L I D C R D * R G N C T E D H * R N K -
 b S C L I E E I D E A T A Q K I I K E I K -
 c L D C L K R L T R Q L H R R S L K K * S -

1501 GTAACCTTTTATTAATTAGAGAATAAACTAAATTAATAATAGAGATCAGCGATCTTCAA 1560
 CATTGAAAAATAATTAATCTCTTATTTGATTAAATGATTATATCTCTAGTCGCTAGAAGTT

a V T F I N * R I N * I T N I E I S D L Q -
 b * L L L I R E * T K L L I * R S A I F N -
 c N F Y * L E N K L N Y * Y R D Q R S S I -

1561 TTGACGAAATAAAAGCTGAACTAAAGTTAGACAATAAAAAATACAAACCTTGGTCAAAAT 1620
 AACTGCTTTATTTTCGACTTGATTCAATCTGTTATTTTTATGTTTGAACCACTTTTA

a L T K * K L N * S * T I K N T N L G Q N -
 b C R N K S C T K V R Q * K I Q T L V K I -
 c D E I K A E L K L D N K K Y K P W S K Y -

1621 ATTGAGGAAGGAAAAGAAGACCAGTTAGCAAAAGAAAAATAAGGCAATAAATAAAATGA 1680
 TAACTCCTTCTTTTCTCTGGTCAATCGTTTCTTTTATTCGGTTATTTATTTTACT

a I E E G K E D Q L A K E K I R Q * I K C -
 b L R K E K K T S * Q K K K * G N K * N E -
 c C G R K R R P V S K R K N K A I N K M S -

[illegible]

a V Q K C R E N K R F I F F S I N L L K R G V -
b Y R E S V E K I * K D L Y F F F S I * I F Y I C E K R E G F -
c T R E S V K K * K K I Y F F F S Q * I F Y I C E K R E G F -

```

a      L  G  F  W  G  F  G  -
b      W  G  F  G  V  L  G  -
c      G  V  L  G  F  W  -

```


FIGURE 14

132 LSTQKQYFFQDEWNQVRAMIGNEL.FRHLYTKYLIFORTSE..GTLVQFC 178
 1 MSRRNQ.....KKPQAPIGNETNLDFVLQNLLEVYKSQIEHYKTQQQI 43
 179 GNNVFDHLKVNDKFDKKQKGAADMNEPRCCSTCKYNVKNKDHFLNNIN 228
 44 KEEDLKLLKFNQDQDGNNGNDDDEE.....NNSNKQQLLRRVN 84
 229 VPNWNNMKSRTTRIFYCTHFNRRNQFFKKHEFVSNKNNISAMDRAQTIFTN 278
 85QIKQOVQLIKK...VGSKEVDLNLNEDENKKN 114
 279 IFRFNRIKKLKDKVIEKIAYMLEKVKDFNFNYLTCKSCPLPENWRERKQ 328
 115 GLSEQQVKEEQRLTITEEQVKYQNLVFNMDYQLDLNESGCHRRHRRRETDY 164
 329 KIENLINKTREEKSKYYEELFSYTTDNKCVTQFINE.FFYNILPKDFLTG 377
 165 DTEKWFEISHDQK.....NYVSIYANQKTSYCWWLKDYFNK 200
 378 RNRKNFQKKVKYVELNKHLEIHNKLLLEKINTREISWMQVETSAXHFY 427
 201 NNYDHLNVSINRLE..TEAEFYAFDDFSQTIKLTNNSYQTVNID..... 242
 428 FDHENIYVLWKLRLWI..FEDLVSLIRCFYVTEQQKSYSTYYRKN 475
 243 VNFDDNLCILALLRFLSLERFNILNIRSSY..TRNQYNFEKIGELLETI 290
 476 WDVIMKMSIADLKKETLAEVQKEVEEWKSLGFPAGKLRLLPKKTTFRP 525
 291 FAVVFSHR.....HLQGIHLQVPCEAFQYLVNSSSQISVKDSQLQ 330
 526 IMTFNKKIVNSDRKTKLTNTKLLNSHMLKTLKNRMFKDPFGFAVFNY 575
 331 VVSFSTDLKLV..TNKVQDYFKLQEFPRLTHVSQAIPVSATNAVENL 378
 576 DDVMKKYEEFVCKWKVQGPQLF.....FATMDIEKCYDS..VNREK 615
 379 NVLLKKVKH..ANLLVSIPTQFNDFYFVNLQHLKLEFGLFEPNLTQK 426
 616 LSTFL.....KTTKLLSSDFWIMTAQILKRKNNI..VIDSKNFRKKEMK 657
 427 LENLLLSIKQSNLKLFLRLNFYTYVAQETSRKQILKQATTIKNLKNNKQ 476
 658 DYFRQKFQKIALEGGOYPTLFSVLEN..EQNDLNAKKTLLIVEAKQRNYFK 705
 477 EETPETKDETPSESTSGMKFFDHLSELTELEDFSVN....LQATQEIY.. 520
 706 KDNELQPVINICQYNYINFNGKFYKOTKGIPOGLCVSSILSSFYATLEE 755
 521 DSLHKLLIRSTNLKKFKLSYKYEKSKMDTFIDLKNI.....YETLNN 564
 756 SSLGLRDESMNPENPNVLLMRLTDDYLLITTOENNAVLFIKLINVSR 805
 565LKRCSVNISNPHGNISYELTNKDSTFYKFKLTNLQE 600
 806 ENGFKFNMKQLQTSFPLSPSKFAKYGMDSVEEQNIVQDYCDWIGISIDMK 855
 601 LQHAKYTFK..QNEFQFNNVKSAKIESSSLESLEDIDSLCKSIASCKNLQ 648
 856 TLALMPNINLRIEGLCTNLNMQT..KKASMWLKK..KLKSFLMNNITH 901
 649 NVNI.....IASLLYPNNIQKNPFKNPNNLLFFKQFEQLKNLENVSINC 691
 902 YFRKTI...TTEDFANKTLNKLFISSGYKMQCAKEYKDHFKKNLAMSSM 948
 692 ILDQHILNSISEFLEKNKKIKAFILKRYLLQYVLDYTKLFTLQQLPEL 741
 949 IDLEVSKIISVT.....RAFFKYLVCNIKDT..IFGEEHY 982
 742 NQYVINQLEELTVSEVHKQVWENHKQKAFYEPLCEFIKESQTLQLIDF 791
 983 PDFFLS..TLKHFIEIFSTKKY..IFNRVCMILKAKEAKLKSQDCQSLIQ 1028
 792 DQNTVSDDSIKKILESISESKYHHYLRNPSQSSSLIKSENEEIQELLK 840

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted November 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Parameter	Unit	Value	Standard Error	t-Statistic	p-Value
Intercept		1.0000	0.0000	1.0000	0.0000
Age	Years	0.0000	0.0000	0.0000	0.0000
Age ²	Years ²	0.0000	0.0000	0.0000	0.0000
Age ³	Years ³	0.0000	0.0000	0.0000	0.0000
Age ⁴	Years ⁴	0.0000	0.0000	0.0000	0.0000
Age ⁵	Years ⁵	0.0000	0.0000	0.0000	0.0000
Age ⁶	Years ⁶	0.0000	0.0000	0.0000	0.0000
Age ⁷	Years ⁷	0.0000	0.0000	0.0000	0.0000
Age ⁸	Years ⁸	0.0000	0.0000	0.0000	0.0000
Age ⁹	Years ⁹	0.0000	0.0000	0.0000	0.0000
Age ¹⁰	Years ¹⁰	0.0000	0.0000	0.0000	0.0000
Age ¹¹	Years ¹¹	0.0000	0.0000	0.0000	0.0000
Age ¹²	Years ¹²	0.0000	0.0000	0.0000	0.0000
Age ¹³	Years ¹³	0.0000	0.0000	0.0000	0.0000
Age ¹⁴	Years ¹⁴	0.0000	0.0000	0.0000	0.0000
Age ¹⁵	Years ¹⁵	0.0000	0.0000	0.0000	0.0000
Age ¹⁶	Years ¹⁶	0.0000	0.0000	0.0000	0.0000
Age ¹⁷	Years ¹⁷	0.0000	0.0000	0.0000	0.0000
Age ¹⁸	Years ¹⁸	0.0000	0.0000	0.0000	0.0000
Age ¹⁹	Years ¹⁹	0.0000	0.0000	0.0000	0.0000
Age ²⁰	Years ²⁰	0.0000	0.0000	0.0000	0.0000
Age ²¹	Years ²¹	0.0000	0.0000	0.0000	0.0000
Age ²²	Years ²²	0.0000	0.0000	0.0000	0.0000
Age ²³	Years ²³	0.0000	0.0000	0.0000	0.0000
Age ²⁴	Years ²⁴	0.0000	0.0000	0.0000	0.0000
Age ²⁵	Years ²⁵	0.0000	0.0000	0.0000	0.0000
Age ²⁶	Years ²⁶	0.0000	0.0000	0.0000	0.0000
Age ²⁷	Years ²⁷	0.0000	0.0000	0.0000	0.0000
Age ²⁸	Years ²⁸	0.0000	0.0000	0.0000	0.0000
Age ²⁹	Years ²⁹	0.0000	0.0000	0.0000	0.0000
Age ³⁰	Years ³⁰	0.0000	0.0000	0.0000	0.0000
Age ³¹	Years ³¹	0.0000	0.0000	0.0000	0.0000
Age ³²	Years ³²	0.0000	0.0000	0.0000	0.0000
Age ³³	Years ³³	0.0000	0.0000	0.0000	0.0000
Age ³⁴	Years ³⁴	0.0000	0.0000	0.0000	0.0000
Age ³⁵	Years ³⁵	0.0000	0.0000	0.0000	0.0000
Age ³⁶	Years ³⁶	0.0000	0.0000	0.0000	0.0000
Age ³⁷	Years ³⁷	0.0000	0.0000	0.0000	0.0000
Age ³⁸	Years ³⁸	0.0000	0.0000	0.0000	0.0000
Age ³⁹	Years ³⁹	0.0000	0.0000	0.0000	0.0000
Age ⁴⁰	Years ⁴⁰	0.0000	0.0000	0.0000	0.0000
Age ⁴¹	Years ⁴¹	0.0000	0.0000	0.0000	0.0000
Age ⁴²	Years ⁴²	0.0000	0.0000	0.0000	0.0000
Age ⁴³	Years ⁴³	0.0000	0.0000	0.0000	0.0000
Age ⁴⁴	Years ⁴⁴	0.0000	0.0000	0.0000	0.0000
Age ⁴⁵	Years ⁴⁵	0.0000	0.0000	0.0000	0.0000
Age ⁴⁶	Years ⁴⁶	0.0000	0.0000	0.0000	0.0000
Age ⁴⁷	Years ⁴⁷	0.0000	0.0000	0.0000	0.0000
Age ⁴⁸	Years ⁴⁸	0.0000	0.0000	0.0000	0.0000
Age					

— 

FIGURE 16

1 MEMDIDLDDIENL.....LPNTFNKYSSSCSDKKGCKTLKSGSKSPS... 42
|:|... || ...|... |:| |...: . | | ...:
491 IELAIAKIAVNKNLDEIKGHTAIFSDVSGSMSTMSGGAKKYGSVRTCLEC 540
43 ..LTIPKLQKQ.....LEFYPSDANLYNDSFLRKLVLKSGEQRVEIETLL 85
|:| : || : : |... :|...| : |...| |...:| |
541 ALVLGLMVKQRCEKSSFYIFSSPSSQCNCYL.EVDLPGDELRPMSMQKLL 589

MEMDIDLDDIENL.....LPNTFNKYSSSCSDKKGCKTLKSGSKSPS... 42
|:|... || ...|... |:| |...: . | | ...:
491 IELAIAKIAVNKNLDEIKGHTAIFSDVSGSMSTMSGGAKKYGSVRTCLEC 540
43 ..LTIPKLQKQ.....LEFYPSDANLYNDSFLRKLVLKSGEQRVEIETLL 85
|:| : || : : |... :|...| : |...| |...:| |
541 ALVLGLMVKQRCEKSSFYIFSSPSSQCNCYL.EVDLPGDELRPMSMQKLL 589

FIGURE 17

	Motif A	Motif B
Consensus	h--h DD h--h--h	h---+ QC --- SP
telomerase p123	GQPKLFFATMDIEKCYDSVNREKLSFLKTTKLL-100-KFYKQTKGIF QC LCVSSILSSFFYYATLEESSLGFL	
Dong (LINE)	KNRNHLHCTYIDYKKAEDSIPHSWLIQVLEIYKIN-28-RQLAIKKGIV QC DSLSPLWFCALNPLSHQLHNDR	
al S.c.(groupII)	FGGSNWFFREVDLKKCFDTISHDLIIKELKRYISD-26-HVPVGPFRVCV QC CAPTSPALCNAVLRLDRRLAGLA	
HIV-RT	LKKKKSVTVLGVGDAYFSVPLDEDFRKYTAFTIP-7-GIRYQYNVLF QC WKGSFAIFQSSMTKILEPFRKQN	
L8543.12 <i>yeast</i>	VLPELYFMKFDVKSCTDSIPMECMRILKDALKN-68-KCYIREDGLE QC ESSLSA?IVDLVYDDLLLEFYSEPK	

	Motif C	Motif D	Motif E
Consensus	h--Y DD h <hh< td=""><td>Ch-h---K</td><td>h-hCh-h</td></hh<>	Ch -h--- K	h-h Ch -h
telomerase p123	-14-LMRLL DD YLLITTTQENN-0-AVLFIKELINVSREN CFKFN TKLQT-23-QDYCDWIGISI		
Dong (LINE)	-16-HLIY DD IKLYAKNDKE-0-MKKLIDTTTIFSNDISMQFGLDKCKT-25-KC-YKYLGFQQ		
al S.c.(groupII)	-55-YVRY DD ILIGVLGSKN-2-KIIKRDNNFLNS IGLT INEERLLI-4-ETPARFLGYNI		
HIV-RT	-4-IYQY DD LVVGSHLEIG-1-HRTKIEELRQHLRLW GLT TPD RR HQK-0-EP?FLMWGYEL		
L8543.12 <i>yeast</i>	-8-ILKL DD FLIISTDQQQ.....VINIKKLAM CF QKYN RR ANR-41-IRSKSSKGIFR		

FIGURE 18

telomerase p43 LQKQLEFYFSDANLYNDSFLRKLVLRSGEQRVETETLLM
human La ICHQVEYFYEGDFNLPRDKFLKEQI.KLDEGWVPLEIMIK
Xenopus LaA ICEQIEYFYGDHNLPRDKFLKQOI.LLDDGWVPLETIK
Drosophila La ILRQVEYFYGDANLNRDKFLREQIGKNEDGWVPLSVLVT
S. c. Lhp1p CLKQVEFYFSEFNFPYDRFLRTTAEK.NDGWVPISTIAT

FIGURE 19

1 aactcattta attactaatt taatcaacaa gattgataaa aagcagtaaa taaaacccaa
 61 tagatttaatt ttgaaaagta tcaattgaaa aatggaaatt gaaaacaact aagcacaata
 121 gccaaaagcc gaaaaattgt ggtgggaact tgaattagag atgcaagaaa accaaaatga
 181 tatataagtt agggttaaga ttgacgatcc taagcaatat ctctgtaacg tcaactgcagc
 241 atgtttgttg taggaaggta gttactacta agataaagat gaaagaagat atatcatcac
 301 taaagcactt ctgaggtgg ctgagtcga tctgagttc atctgctagt tggcagtcta
 361 catccgtaat gaactttaca tcagaactac cactaactac attgtagcat ttgtgttgt
 421 ccacaagaat actcaacat tcacgaaaa gtacttcaac aaagcagtac tttgcctaa
 481 tgacttactg gaagtctgtg aattgcata ggttctctat attttgatg caactgaatt
 541 caaaaatttg tatcttgata ggatacttc ataagatatt cgtaagggaac tcactttccg
 601 taagtgttta caaagatgcg tcagaagcaa gtttctgaa ttcaacgaat actaacttgg
 661 taagtattgc actgaatcct aacgtaagaa aacaatgttc cgttacctct cagtaccaa
 721 caagtaaaag tgggattaaa ctaagaagaa gagaaaagag aatctcttaa ccaaacttta
 781 ggcaataaag gaatctgaag ataagtcaa gagagaaact ggagacataa tgaacgttga
 841 agatgcaatc aaggctttaa aaccagcagt tatgaagaaa atagccaaga gatagaatgc
 901 catgaagaaa cacatgaagg cacctaaaat tcttaactct accttggaaat caaagtactt
 961 gaacttcaag gatctcatta agttctgcca tattctgag cctaaagaaa gagtctataa
 1021 gatccttggt aaaaaatacc ctaagaccga agaggaatac aaagcagcct ttggtgattc
 1081 tgcatctgca ccttcaatc ctgaattggc tggaaagcgt atgaagattg aaatctctaa
 1141 aacatgggaa aatgaactca gtgcaaaagg caacactgct gaggtttggg ataatttaat
 1201 ttcaagcaat taactcccat atatggccat gttacgtaac ttgtctaaca tcttaaaagc
 1261 cgggtttica gatactacac actctattgt gatcaacaag atttgtgagc ccaaggccgt
 1321 tgagaactcc aagatgttcc ctcttcaatt cttagtgcc attgaagctg ttaatgaagc
 1381 agttactaag ggattcaagg ccaagaagag agaaaatag aatcttaaag gtcaaatcga
 1441 agcagtaaaag gaagtgttg aaaaaaccga tgaagagaag aaagatatgg agttggagta
 1501 aaccgaagaa ggagaatttg ttaaagcaa cgaaggaatt ggcaagcaat acattaactc
 1561 cattgaactt gcaatcaaga tagcagttaa caagaattta gatgaaatca aaggacacac
 1621 tgcaatcttc tctgatgttt ctggttctat gactacctca atgtcagggt gagccaagaa
 1681 gtatggttcc gttcgtactt gtctcgagtg tgcattagtc ctggtttga tggtaaaata
 1741 acgttgtgaa aagtcctcat tctacatctt cagttcacct agttctcaat gcaataagt
 1801 ttacttagaa gttgatctcc ctggagacga actccgtctt tctatgtaaa aacttttgca
 1861 agagaagga aaacttggtg gtggtactga ttccctctat gagtgcattg atgaatggac
 1921 aaagaataaa actcacgtag acaatatcgt tattttgtct gatatgatga ttgcagaagg
 1981 atattcagat atcaatgtta gaggcagttc cattgttaac agcatcaaaa agtacaagga
 2041 tgaagtaaat ctaacatta aaatctttgc agttgactta gaaggttacg gaaagtcct
 2101 taatctaggt gatgagtca atgaaaacaa ctacatcaag atattcggtg tgagcgattc
 2161 aatcttaaag ttcatttcag ccaagcaagg aggagcaaat atggtcgaag ttatcaaaaa
 2221 ctttgcctt caaaaaatag gacaaaagt agtttctga gattcttcta taacaaaaat
 2281 ctacccccac tttttgtt tattgcatag ccattatgaa atttaatta ttatctatt
 2341 atttaagtta ctacatagt ttatgtatcg cagtctatta gcctattcaa atgattctgc
 2401 aaagaacaaa aaagattaaa a

FIGURE 20

MEIENNQAQQPKAEKLWWELELEMQENQNDIQVRVKIDDPKQYL
VNVTAACLLQEGSYQDKDERRYITKALLEVAESDPEFICQLAVYIRNELYIRTTN
YIVAFCVVHKNTQPFIEKYFNKAVLLPNDLLEVCEFAQVLYIFDATEFKNLYLDRILS
QDIRKELTFRKCLQRCVRSKFSEFNEYQLGKYCTESQRKKTMFRYLSVTNKQKWDQTK
KKRKENLLTKLQAIKESEDKSKRETGDIMNVEDAIKALKPAVMKKIAKRQONAMKKHMK
APKIPNSTLESKYLTFKDLIKFCHISEPKERVYKILGKKYPKTEEEYKAAFGDSASAP
FNPELAGKRMKIEISKTWENELSAKGNTAEVWDNLISSNQLPYMAMLRNLSNILKAGV
SDTTHSIVINKICEPKAVENSKMFPLQFFSAIEAVNEAVTKGFKAKKRENMNLKGQIE
AVKEVVEKTDEEKKDMELEQTEEGEFVKVNEGIGKQYINSIELAIKIAVNKNLDEIKG
HTAIFSDVSGSMSTMSGGAKKYGSVRTCLECALVLGLMVKQRCEKSSFYIFSSPSSQ
CNKCYLEVDLPGDEL RPSMQKLLQEKGKLGGGTDFPYECIDEWTKNKTHTVDNIVILSD
MMIAEGYSDINVRGSSIVNSIKKYKDEVNPNIKIFAVDLEGY GKCLNLGDEFNENNYI
KIFGMSDSILKFISAKQGGANMVEVIKNFALQKIGQK

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 21

1 tcaatactat taattaataa ataaaaaaaa gcaaaactaca aagaaaatgt caaggcgtaa
 61 ctaaaaaaag ccataggctc ctataggcaa tgaacaaat ctgattttg tattacaaaa
 121 tctagaagtt tacaaaagcc agattgagca ttataagacc tagtagtaat agatcaaaga
 181 ggaggatctc aagcttttaa agttcaaaaa ttaagattag gatggaaact ctggcaacga
 241 tgatgatgat gaagaaaaca actcaataa ataataagaa ttattaagga gagtcaatta
 301 gattaagtag caagttaat tgataaaaa agttggttct aaggtagaga aagatttgaa
 361 ttgaacgaa gatgaaaaca aaaagaatgg actttctgaa tagcaagtga aagaagagta
 421 attaagaacg attactgaag aatagggtta gtattaaaat ttagtattta acatggacta
 481 ccagttagat ttaaatgaga gtggtggcca tagaagacac agaagagaaa cagattatga
 541 tactgaaaaa tggtttgaaa tatctcatga ccaaaaaaat tatgtatcaa ttacgccaa
 601 ctaaaagaca tcatattgtt ggtggcttaa agattatttt aataaaaaaca attatgatca
 661 tcttaatgta agcattaaca gactagaaac tgaagccgaa ttctatgcct ttgatgattt
 721 ttcacaaaca atcaaaacta ctaataattc ttactagact gttacatag acgttaattt
 781 tgataataat ctctgtatc tcgactgtct tagattttta ttactactag aaagattcaa
 841 tattttgaat ataagatctt ctatacaag aaattaatat aattttgaga aaattggtga
 901 gctacttgaa actatcttcg cagttgtctt ttctcatcgc cactacaag gcattcattt
 961 acaagttcct tgcgaagcgt tctaatattt agttaactcc tcatcataaa ttacggttaa
 1021 agatagctaa ttataggtat actctttctc tacagactta aaattagttg acactaaca
 1081 agtccaagat tattttaagt tcttataaga attccctcgt ttgactcatg taagctagta
 1141 ggctatccca gttagtgcta ctaacgctgt agagaacctc aatgttttac ttaaaaagg
 1201 caagcatgct aatcttaatt tagtttctat ccctacctaa ttcaattttg atttctactt
 1261 tgtaattta taacatttga aattagagtt tggattagaa ccaaatattt tgacaaaaca
 1321 aaagcttgaa aatctacttt tgagtataaa ataatacaaa aatcttaaat ttttaagatt
 1381 aaacttttac acctacgttg cttagaagac ctccagaaaa cagatattaa aacaagctac
 1441 aacaatcaaa aatctcaaaa acaataaaaa tcaagaagaa actcctgaaa ctaagatga
 1501 aactccaagc gaaagcacaa gtggtatgaa atttttgat catctttctg aattaaccga
 1561 gcttgaagat ttacggtta acttgaagc tacccaagaa atttatgata gcttgcacaa
 1621 acttttgatt agatcaacaa atttaagaa gttcaaatta agttacaat atgaatgga
 1681 aaagagtaaa atggatacat tcatagatct taagaatatt tatgaacct taaacaatct
 1741 taaaagatgc tctgttaata tatcaaatcc tcatggaaac atttctatg aactgacaaa
 1801 taaagattct acttttata aatttaagct gaccttaac taagaattat aacacgctaa
 1861 gtatactttt aagtagaacg aatttaatt taataacgtt aaaagtgc aaattgaatc
 1921 ttctcatta gaaagcttag aagatattga tagtctttgc aaatctattg cttctgttaa
 1981 aaatttaca aatgttaata ttatcgccag ttgctctat cccaacaata tttagaaaa
 2041 tctttcaat aagcccaatc ttctattttt caagcaattt gaataattga aaaatttga
 2101 aaatgtatct atcaactgta ttctgatca gcatatactt aattctattt cagaattctt
 2161 agaaaagaat aaaaaataa aagcattcat ttgaaaaga tattatttat tacaatatta
 2221 tcttgattat actaaattat taaaacact tcaatagtta cctgaattaa attaagtta
 2281 cattaattag caattagaag aattgactgt gagtgaagta cataagtaag tatgggaaaa
 2341 ccacaagcaa aagctttct atgaaccatt atgtgagttt atcaaagaat catcctaaac
 2401 cctttagcta atagattttg accaaaacac tgaagtgtat gactctatta aaaagatttt
 2461 agaactata tctgagtcta agtatcatca ttatttgaga ttgaacccta gttaatctag
 2521 cagtttaatt aaatctgaaa acgaagaat ttaagaactt ctcaaagctt gcgacgaaaa
 2581 aggtgtttta gtaaaagcat actataaatt cctctatgt ttaccaactg gtacttatta
 2641 cgattacaat tcagatagat ggtgattaat taaatattag tttaataaa tattaatat
 2701 tgaattttc ttgcttatt atttgaataa tacatacaat agtcattttt agtgttttga
 2761 atatatttta gttatttaatt tcattatttt aagtaaataa ttatttttca atcattttt
 2821 aaaaaatcg

FIGURE 22

MSRRNQKKPQAPIGNETNLDVFLQNLEVYKSQIEHYKTQQQQIK
EEDLKLLKFKNQDQDGNsgnddddeennsnkqqellrrvnqikqqvqlikkvgskvek
DLNLNEDENKKNGLSEQQVKEEQLRTITEEQVKYQNLVFNMDYQLDLNESGGHRRHRR
ETDYDTEKWFEISHDQKNYVSIYANQKTSYCWWLKDYFNKNNYDHLNVSINRLETEAE
FYAFDDFSQTIKLTNNSYQTVNIDVNFNNLCILALLRFLSLERFNILNIRSSYTRN
QYNFEKIGELLETIFAVVFSHRHLQGIHLQVPCEAFQYLVNSSSQISVKDSQLQVYSF
STDLKLVD TNKVQDYFKFLQEFPR LTHVSQQAIPVSATNAVENLNVLLKKVKHANLNL
VSIPTQFNFDYFVNLQHLKLEFGLEPNILTKQKLENLLLSIKQSKNLKFLRLNFYTY
VAQETSRKQILKQATTIKNLKNNKNQEETPETKDETPSESTSGMKFFDHLSELTELED
FSVNLQATQEIYDSLHKL LIRSTNLKKFKLSYKYEMEKSKMDTFIDLKNIYETLNNLK
RCSVNISNPHGNISYELTNKDSTFYKFKLTLNQELQHAKYTFKQNEFQFNNVKSAKIE
SSSLESLEDIDSLCKSIASCKNLQNVNIIASLLYPNNIQKNPFNKPNLLFFKQFEQLK
NLENVSINCILDQHILNSISEFLEKNKKIKAFILKRYYYLLQYYLDYTKLFKTLQQLPE
LNQVYINQQLEELTVSEVHKQVWENHKQKAFYEPLCEFIKESSQTLQLIDFDQNTVSD
DSIKKILESISESKYHHYLRLNPSQSSSLIKSENEEIQELLKACDEKGVLVKAYYKFP
LCLPTGTYDYDYNDRW

FIGURE 23

MKILFEFIQDKLDIDLQTNSTYKENLKCGHFNGLDEILTTCFAL
PNSRKIALPCLPGDLSHKAVIDHCIIYLLTGELYNNVLTFGYKIARNEDVNNSLFCHS
ANVNVTLKGAAWKMFHSLVGTYAFVDLLINYTVIQFNGQFFTQIVGNRCNEPHLPPK
WVQRSSSSSATAAQIKQLTEPVTNKQFLHKLNNSSSFFPYSKILPSSSSIKKLTDLR
EAIPTNLVKIPQRLKVRJNLTLQKLLKRHKRLNYVSILNSICPPLEGTVLDLSHLR
QSPKERVLFIIQKLLPQEMFGSKKNKGKIIKNLNLLSLPLNGYLPFDSLLKKL
RLKDFRWLFISDIWFTKHNFNENLNQLAICFISWLFRLIPKIIQTFFYCTEISSTVTI
VYFRHDTWNKLITPFIVEYFKTYLVENNVCRNHNSYTLNHNHSMRIIPKKSNNFR
IIAIPCRGADEEEFTTYKENHKNAIQPTQKILEYLRNKRPTSFTKIYSPTQIADRIKE
FKQRLKKFNNVLPELYFMKFDVKSCYDSIPRMECMRILKDALKNENGFFVRSQYFFN
TNTGVLKLFNVVNASRVPKPYELYIDNVRTVHLSNQDVINVVEMEIFKTALWVEDKCY
IREDGLFQGSSLSAPIVDLVYDDLLEFYSEFKASPSQDTLILKLADDFLIISTDQQQV
INIKKLAMGGFQKYNAKANRDKILAVSSQSDDDTVIQFCAMHIFVKELEVWKHSSTMN
NFHIRSKSSKGIFRSLIALFNTRISYKTIDTNLNSTNTVLMQIDHVVKNISECYKSAF
KDLSINVTQNMQFHSFLQRIEMTVSGCPITKCDPLIEYEVRFITLNGFLESLSNTS
KFKDNIILLRKEIQHLQAYIYIYIHVN

FIGURE 24

<i>Oxytricha</i>	LCVSYILSSFYYANLEENALQFLRKESMDPEKPETNLLMRLT
<i>Euplotes</i>	LCVSSILSSFYYATLEESSLGFLRDESMNPENPNVNLLMRLT

FIGURE 25

Motif 0

human	AKFLHWLMSVYVVELLRSFFYVTETTFQKNR
tez1	ISEIEWLVLGKRSNAKMCLSDFEKRRQIFAEFIYWLNSFIIPILQSFFYITESSDLNR
EST2	LKDPRWLFISD--IWFTKHNFNENLNQLAICFISWLFRLIPKIIQTFFYCTEISSTVT-
p123	TREISWMQVET-SAKHPYFDHEN-IYVLWKLLRWIFEDLVVSLIRCFFYVTEQQKSYSK
	. * * * *

Motif 1

human	LFFYRKSVWSKLSIGIRQHLKRVQLRDVSEAEVRQHRARPALLTSRLRFIPKP--DGL
tez1	TVYFRKDIWKLLCRPFI-TSMKMEAFEKINENNVRMDTQK-TTLPPAVIRLLPKK--MTF
EST2	IVYFRHDTWNKLITPFIVEYFKTYLVENNVCRNHNSTLS--NFNHSMRIIPKKSNEF
p123	TYYYRKNIWDVIMKMSI-ADLKETLAEVQEKEVEENKKS-LGFAPGKLRLIPKK--TTF
	..* . * . * . * . * . *

Motif 2

human	RPIVNM DYVVGARTFRREKRAERLTSRVKALF-SVLNYERA
tez1	RLITN-LRKRFLIKMGSNKKMLVSTNQTLRPVASILKHLNEESSGIPFNLEVYMKLLTF
EST2	RIIAIPCRGADEEEFTIYKENHKNAIQPTQKILEYLRNKRPTSFTKIYSPTQIADRIKEF
p123	RPIMTFNKKIVNSDRKTTKLTNTKLLNSHMLKTLKN-RMFKDPPGFAVFNYDDVMKKY
	* * . * . *

Motif 3 (A)

tez1	KKDLLKHRMFGK-RKYFVRIDIKSCYDRIKQDLMFRIVKK-KLKDPEFVIRKYATIHATS
EST2	RQRLKKKFNNVLPELYFMKFDVKSCYDSIPRMECMRILKD-ALKNENGFFVRSQYFFNTN
p123	EEFVCKWKQVGQPKLPFATMDIEKCYDSVNRKELSTFLKTTKLLSSDFWIMTAQILKRKN
	. * . * . * * * . * *

FIGURE 26

ATTTATACTCATGAAAATCTTATTCGAGTTCATTCAAGACAAGCTTGACATTGATCTACA
 GACCAACAGTACTTACAAAGAAAATTTAAAATGTGGTCACTTCAATGGCCTCGATGAAAT
 TCTAACTACGTGTTTCGCACTACCAAATTCAGAAAAATAGCATTACCATGCCTTCCTGG
 TGACTIONAAGCCACAAAGCAGTCATTGATCACTGCATCATTTACCTGTTGACGGGCGAATT
 ATACAACAACGTACTAACATTTGGCTATAAAATAGCTAGAAATGAAGATGTCAACAATAG
 TCTTTTTTTGCCATTCTGCAAATGTTAACGTTACGTTACTGAAAGGCGCTGCTTGGAAAAT
 GTTCCACAGTTTGGTCGGTACATACGCATTCGTTGATTTATTGATCAATTATACAGTAAT
 TCAATTTAATGGGCAGTTTTTCACTCAAATCGTGGGTAACAGATGTAACGAACCTCATCT
 GCCGCCCAAATGGGTCCAACGATCATCCTCATCATCCGCAACTGCTGCGCAAATCAAACA
 ACTTACAGAACCAGTGACAAATAACAATTCTTACACAAGCTCAATATAAATTCCTCTTC
 TTTTTTTCCTTATAGCAAGATCCTTCCTTCATCATCATCTATCAAAAAGCTAACTGACTT
 GAGAGAAGCTATTTTTCCCACAAATTTGCTTAAATTCCTCAGAGACTAAAGGTACGAAT
 TAATTTGACGCTGCAAAAGCTATTTAAAGAGACATAAGCGTTTGAATTACGTTTCTATTTT
 GAATAGTATTTGCCACCATTGGAAGGGACCGTATTGGACTTGTGCGCATTTGAGTAGGCA
 ATCACCAAAGGAACGAGTCTTGAAATTTATCATTGTTATTTTACAGAAGTTATTACCCCA
 AGAAATGTTTGGCTCAAAGAAAAATAAAGGAAAAATTATCAAGAATCTAAATCTTTTATT
 AAGTTTACCCTTAAATGGCTATTTACCATTTTGATAGTTTGTGAAAAAGTTAAGATTAA
 GGATTTTCGGTGGTTGTTTCATTTCTGATATTTGGTTTACCAAGCACAAATTTTGAAACTT
 GAATCAATTGGCGATTTGTTTCATTTCTGGCTATTTAGACAATAATTCCCAAAATTAT
 ACAGACTTTTTTTTACTGCACCGAAATATCTTCTACAGTGACAATTGTTTACTTTAGACA
 TGATACTTGGAATAAACTTATCACCCCTTTTATCGTAGAATATTTTAAGACGTACTTAGT
 CGAAAACAACGTATGTAGAAACCATAATAGTTACACGTTGTCCAATTTCAATCATAGCAA
 AATGAGGATTATACCAAAAAAAGTAATAATGAGTTTCAAGATTATTGCCATCCCATGCAG
 AGGGGCAGACGAAGAAGAATTCACAATTTATAAGGAGAATCACAAAAATGCTATCCAGCC
 CACTCAAAAAATTTTAGAATACCTAAGAAACAAAAGGCCGACTAGTTTTACTAAAATATA
 TTCTCCAACGCAATAGCTGACCGTATCAAAGAATTTAAGCAGAGACTTTTAAAGAAATT
 TAATAATGTCTTACCAGAGCTTTATTTTCATGAAATTTGATGTCAAATCTTGCTATGATTC
 CATACCAAGGATGGAATGTATGAGGATACTCAAGGATGCGCTAAAAAATGAAATGGGT
 TTTTCGTTAGATCTCAATATTTCTTCAATACCAATACAGGTGTATTGAAGTTATTTAATGT
 TGTTAACGCTAGCAGAGTACCAAAACCTTATGAGCTATACATAGATAATGTGAGGACGGT
 TCATTTATCAAATCAGGATGTTATAAACGTTGTAGAGATGGAAATATTTAAACAGCTTT
 GTGGGTTGAAGATAAGTGCTACATTAGAGAAGATGGTCTTTTTTCAGGGCTCTAGTTTATC
 TGCTCCGATCGTTGATTTGGTGTATGACGATCTTCTGGAGTTTTATAGCGAGTTTAAAGC
 CAGTCCTAGCCAGGACACATTAATTTTAAACTGGCTGACGATTTCTTATAATATCAAC
 AGACCAACAGCAAGTGATCAATATCAAAAAGCTTGCCATGGGCGGATTTCAAAAATATAA
 TGCGAAAGCCAATAGAGACAAAATTTAGCCGTAAGCTCCAATCAGATGATGATACGGT
 TATTCAATTTTGTGCAATGCACATATTTGTTAAAGAATTGGAAGTTTGGAAACATTCAAG
 CACAATGAATAATTTCCATATCCGTTGAAATCTAGTAAAGGATATTTGGAAGTTTAAAT
 AGCGCTGTTTAACTAGAAATCTCTTATAAAACAATTGACACAAATTTAAATCAACAAA
 CACCGTTCTCATGCAATTTGATCATGTTGTAAAGAACATTTGGAATGTTATAAATCTGC
 TTTTAAGGATCTATCAATTAATGTTACGCAAAATATGCAATTTTCAATCGTTCTTACAACG
 CATCATTGAAATGACAGTCAGCGGTTGTCCAATTACGAAATGTGATCCTTTAATCGAGTA
 TGAGGTACGATTCACCATATTTGAATGGATTTTGGAAAGCCTATCTTCAAACACATCAAA
 ATTTAAAGATAATATCATTCCTTTTGAGAAAGGAAATTCAACACTTGCAAGC

FIGURE 27

AKFLHWLMSVYVVELLSFFYVTETTFQKNRLFFYRKS VWSKLQSIGIRQHLKR
VQLRDVSEAEVRQHREARPAALLTSRLRFIPKPDGLRPIVNM DYVVGARTFRREKR
AERLTSRVKALFSVLNYERA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1

[illegible]

FIGURE 29

MTEHHTPKSRILRFLENQYVYLCTLNDYVQLVLRGSPASSYSNICERLRSDVQTSFSIFLHSTVVGF
 DSKPDEGVQFSSPKCSQSELIANVVKQMFDESFERRRNLLMKGFSMNHEDFRAMHVNGVQNDLV
 STFPNYLISILESKNWQLLLEIGSDAMHYLLSKGSIFEALPNDNYLQISGIPLFKNNVFEETVSKKRK
 RTIETSITQNK SARKEVSWNSISISRFSIFYRSSYKKFKQDL YFNLHSICDRNTVHMWLQWIFPRQFG
 LINAQVQQLHKVIPLVSQSTVVPKRLLLKVYPLIEQTA KRLHRISLSKVYNHYCPYIDTHDDEKILS
 YSLKPNQVFAFLRSILVRVFPKLIWGNQRIFEIILKDLETFLKLSRYESFSLHYLMSNIKISEIEWLVL
 GKRSNAKMCLSDFEK RKQIFAEFIYWL YNSFIIPILQSFFYITESSDLRNRTVYFRKDIWKLLCRPFIT
 SMKMEAFEKINENNVRMDTQKTTLP PAVIRLLPKKNTFRLITNLRKRFLIKMGSNKKMLVSTNQT
 LRPVASILKHLINEESSGIPFNLEVYMKLLTFKKDLLKHRMFGRKKYFVRIDIKSCYDRIKQDLMFR
 IVKKKLLKDPEFVIRKYATI HATSDRATKNFVSEAFSYFDMVPFEKVVQLLSMKTSDTLFVDFVDY
 WTKSSSEIFKMLKEHLSGHIVKIGNSQYLQKVGIPQGSILSSFLCHFYMEDLIDEYLSFTKKKGSVL
 LRVVDDFLFITVNKKDAKKFLNLSLRGFEKHNFST SLEKTVINFENSNGIINNTFFNESKKRMPFFG
 FSVNMRSLDTLLACPKIDEALFNSTSVELTKHMGKSFFYKILRSSLASFAQVFIDITHNSKFNSCCNI
 YRLGYSMCMRAQAYLKRMKDIFIPQRMFITDLLNVIGRKIWKKLAEILGYTSRRFLSSAEVKWLFCL
 LGMRDGLKPSFKYHPCFEQLIYQFQSLTDLIKPLRPVLRQVFLHRRID

FIGURE 30

ggtaccgatttactttcttctcataagctaattgcttctcgaacgctcctaatactctggaatattttacaagaactcaataacaataccaagtcaaattccaatatgaagg
 tgttattagtgatcgataatatttctattttatcggtcgttaccaagtataaggacaaaagaacaactccttccccctaaagacttttacttttataatttacttttcaaatatattcg
 ggttcgcttacttttaactggtgactgttttagctgctacttctagccaaccgctgttttctaccccgctcattggatagctcttggagtagctcacagaaatccttacaatctt
 ctgatgagactatattagattcattacagtcgctgcatattcctaactggagccttacacttttagatgagtcacgctcgcagtgatggagtagtttggtatcatccaacgtttgccttg
 aaaagggtgataattttgcaaatcatgtccttagtggtggaatccgcgaaagtttttgatgcttgcacacgcttagcatgattgagatattcaaaaatttctatccactacaa
 ctcttttaacgcggttttatttttctatttctcatgttgcctcaaatatgtatcatctcgtattagcgtttttccgttttactcctggaatcgtagcttttctactattccccctaatg
 aataatctaaattagtttcgcttataattgatagtagaagaattggtgatttactcgtgtaattgtattagtttaagatactttgcaaaacatttattagctatcattatataaaa
 aaaaactataattataatattaatcaatatttgcggtcactatttattaaacgttatgacagtaggacactttgcatatatatagttatgcttaattggttactgttaacttgcAT
 GACCGAACACCATACCCCCAAAAGCAGGATTCTTCGCTTTCTAGAGAATCAATATGTATACCTATGTA
 CCTTAAATGATTATGTACAACTTGTGAGAGGGTCGCCGGCAAGCTCGTATAGCAATATATGCGAA
 CGCTTGAGAAGCGATGTACAAACGTCCTTTTCTATTTTCTTCATTCTGACTGTAGTCGGCTTCGACAGT
 AAGCCAGATGAAGGTGTTCAATTTCTTCTCCAAAATGCTCACAGTCAGAGGtatatatattttgtttgatttttctattcg
 ggatagctaatatatgggcagCTAATAGCGAATGTTGTAAACAGATGTTTCGATGAAAGTTTGTAGCGTCGAAGGA
 ATCTACTGATGAAAGGGTTTTCCATGgtaagggtatttctaattgtgaaatatttaccgtcaattactgtttcaagagattgttattaaacgataaagAA
 TCATGAAGATTTTCGAGCCATGCATGTAAACGGAGTACAAAATGATCTCGTTTTCTACTTTTCTTAATTA
 CCTTATATCTATACTTGAGTCAAAAAATTGGCACTTTGTAGAAATgtaaataccggttaagatgttgccactttgaaca
 agactgacaagtatGATCGGCAGTGATGCCATGCACTACTTATTATCCAAAGGAAGTATTTTGTAGGCTCTTC
 CAAATGACAATTAACCTTCAGATTTCTGGCATACCACTTTTAAAAATAATGTGTTTGAGGAACTGTGT
 CAAAAAAGAAAGCGAACCATTGAAACATCCATTACTCAAAATAAAAGCGCCCGCAAAGAAGTTTC
 CTGGAATAGCATTTCAATTAGTAGGTTTAGCATTTTTTACAGGTCATCCTATAAGAAGTTTAAGCAAGgt
 aactaactgttatccttcataactaatttttagATCTATATTTTAACTTACACTCTATTTGTGATCGGAACACAGTACACATG
 TGGCTTCAATGGATTTTTTCCAAGGCAATTTGGACTTATAAACGCATTTCAAGTGAAGCAATTGCACAA
 AGTGATTCCACTGGTATCACAGAGTACAGTTGTGCCCAAACGTCCTAAAGGTATACCCTTTAATTGA
 ACAAACAGCAAAGCGACTCCATCGTATTTCTCTATCAAAAGTTTACAACCATTATTGCCCATATATTGA
 CACCCACGATGATGAAAAAATCCTTAGTTATTCCTTAAAGCCGAACCAGGTGTTTGCCTTCTTCGATC
 CATTCTGTTCGAGTGTTTCTAAATTAATCTGGGGTAACCAAAGGATATTTGAGATAATATTAAGg
 tattgtataaaatttattaccactaacgattttaccagACCTCGAAACTTTCTTGAAATTATCGAGATACGAGTCTTTTAGTTTAC
 ATTATTTAATGAGTAACATAAAAGtaaatatgccaattttttaccattaattaacaatcagATTTCAGAAATTGAATGGCTAGT
 CCTTGGAAGAAAGGTCAAATGCGAAATGTGCTTAAGTGATTTTGAGAAACGCAAGCAAATATTTGCGG
 AATTCATCTACTGGCTATACAATTCGTTTATAATACCTATTTTACAATCTTTTTTTTATATCACTGAATC
 AAGTGATTTACGAAATCGAACTGTTTATTTTAGAAAAGATATTTGGAAGCTCTGTGCCGACCCTTTAT
 TACATCAATGAAAATGGAAGCGTTTGAAAAAATAAACGAGgtatttttaagtatttttgcaaaaagctaatttttcagAACAA
 TGTTAGGATGGATACTCAGAAAACACTTTGCCTCCAGCAGTTATTCGTCTATTACCTAAGAAGAATAC
 CTTTCGTCTCATTACGAATTTAAGAAAAAGATTCTTAATAAAAGtattattttgtcatcaatgtactttacttctaacttattattag
 cagATGGGTTCAAACAAAAAATGTTAGTCAGTACGAACCAAACTTTACGACCTGTGGCATCGATACTG
 AAACATTTAATCAATGAAGAAAGTAGTGGTATTCCATTAACTTGGAGGTTTACATGAAGCTTCTTACT
 TTTAAGAAGGATCTTCTTAAGCACCGAATGTTTGGgtaattatataatgcgcgattcctcattattaattttgcagGCGTAAGAAG
 TATTTTGTACGGATAGATATAAAATCCTGTTATGATCGAATAAAGCAAGATTTGATGTTTCGGATTGTT
 AAAAAGAACTCAAGGATCCCGAATTTGTAATTCGAAAGTATGCAACCATACATGCAACAAGTGACCG
 AGCTACAAAAAATTTGTTAGTGAGGCGTTTCTCTATTGtaagtttatttttcattggaatttttaacaaattcttttttagTTGATAT
 GGTGCCTTTGAAGAAAGTCGTGCAGTTACTTTCTATGAAAACATCAGATACTTTGTTTGTGATTGTTTGT
 GGATTATTGGACCAAAAGTTCTTCTGAAATTTTAAAAATGCTCAAGGAACATCTCTCTGGACACATTGT
 TAAAGgtataccaattgttgtaattgaataacactaatgaaactagATAGGAAATTCTCAATACCTTCAAAAAGTTGGTATCCCTC
 AGGGCTCAATTCTGTCTATCTTTTTTGTGTCTTTCTATATGGAAGATTTGATTGATGAATACCTATCGTT
 TACGAAAAAGAAAGGATCAGTGTTGTACGAGTAGTCGACGATTTCTCTTTATAACAGTTAATAAAA
 AGGATGCAAAAAAATTTTTGAATTTATCTTTAAGAAGtgtagttgctgtcattcctaagttctaaccgttgagGATTTGAGAA
 ACACAATTTTCTACGAGCCTGGAGAAAACAGTAATAAACTTTGAAAATAGTAATGGGATAATAAACA
 ATACTTTTTTTAATGAAAGCAAGAAAAGAATGCCATTCTTCGGTTTCTCTGTGAACATGAGGTCTCTTG
 ATACATTGTTAGCATGTCCTAAAATTGATGAAGCCTTATTTAACTCTACATCTGTAGAGCTGACGAAAC
 ATATGGGGAAATCTTTTTTTTACAAAATTCTAAGgtatattgtgaactgaataatagctgacaataatcagATCGAGCCTTGC
 ATCCTTTGCACAAGTATTTATTGACATTACCCACAATTCAAAATTCATTCTTGCTGCAATATATATAG
 GCTAGGATACTCTATGTGTATGAGAGCACAAGCATACTTAAAAAGGATGAAGGATATATTTATCCCC
 AAAGAATGTTTCATAACGGgtgagtacttattttaactgaaaagtcattaattaaccttagATCTTTTGAATGTTATTGGAAGAAAA
 ATTTGAAAAAGTTGGCCGAAATATTAGGATATACGAGTAGGCGTTTCTGTCTCTGCAGAAGTCAA

FIGURE 30 (cont.)

ATGgtacgtgtcggctcgcagacttcagcaatattgacacatcagGCTTTTTTGTCTTGGAATGAGAGATGGTTTGAAACCCTCTT
TCAAATATCATCCATGCTTCGAACAGCTAATATACCAATTTCAAGTCATTGACTGATCTTATCAAGCCGC
TAAGACCAGTTTTGCGACAGGTGTTATTTTTACATAGAAGAATAGCTGATTAAAtgcattttcaattattatatacatcctt
tattactggtgtcttaacaatattactaagtagctgaccccaaagcaagcactataggattttagtaaaagtaaatctcgtattagtttgattgactgtctt
atccttatacttttaagaaagattgacagtgggtgctgactactgccacatgccattaaacgggagtggttaaacattaaaagtaatacatgaggctaattcctttcatttag
aataaggaaagtggttttctataatgaataatgccgcactaatgcaaaaagacgaagattatctttaaacaaggggattaagcatatccgaaggaaaagagagtaatat
accagtggtgttgaaagaaagcaaggataatttgaacaagcttctgcagatgacaggctaaatttggtagccgaatttggtaaaagcccagggtatccatggtggccg
gccttgctactgagacgaaaagaaactaaggatagttgaataactaatagctcattaatgtcttatataagggtttgttttctgacttcaatttgcattgggtgaaaagaaata
gtgttaagccattattgattccgaaatagccaaatttcttggttcctcaaagcgaagtctaaagaacttattgaagcttatgaggcttcaaaaactcctcctgatttaaaggag
gaatctccaccgatgaggaaatggatagcttatcagctgctgaggagaagcctaatttttgcaaaaagaaaatcattgggagacatctcttgatgaatcagatgcgga
gagtatctccagcggatccttgatgtcaataacttctatttctgaaatgtatgtctactgtcgttcgacttctcgtagctctacgcagttaagtgaccaaagggtacc

FIGURE 32

S-1: FFY VTE TTF QKN RLF FYR KSV WSK
S-2: RQH LKR VQL RDV SEA EVR QHR EA
S-3: ART FRR EKR AER LTS RVK ALF SVL NYE

A-1: AKF LHW LMS VYV VEL LRS FFY VTE TTF Q
A-2: LFF YRK SVW SKL QSI GIR QHL KRV QLR DVS
A-3: PAL LTS RLR FIP KPD GLR PIV NMD YVV

THESE ARE THE RESULTS OF THE ANALYSIS

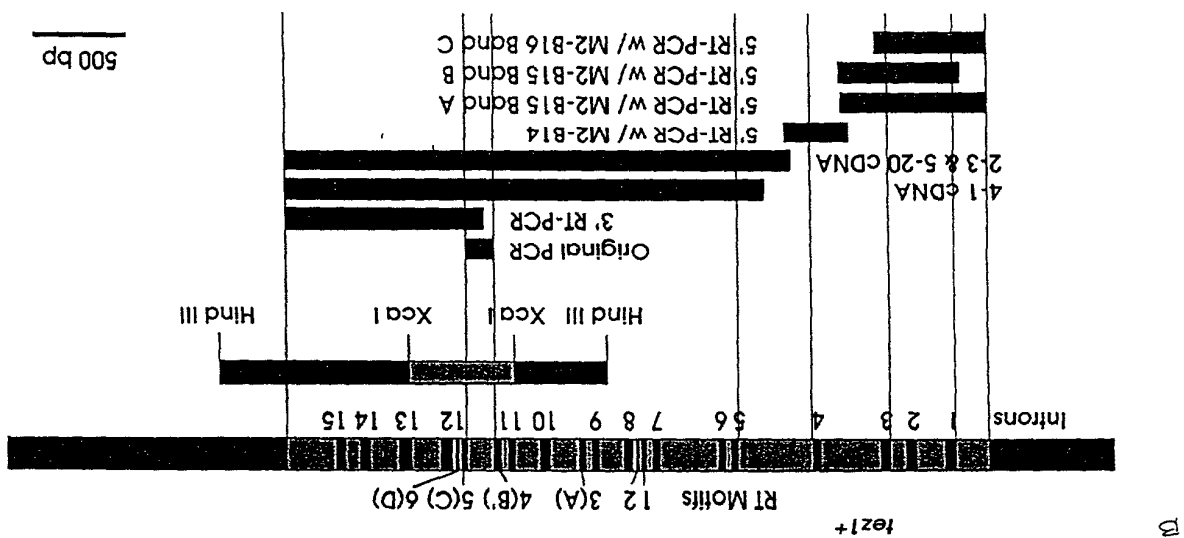
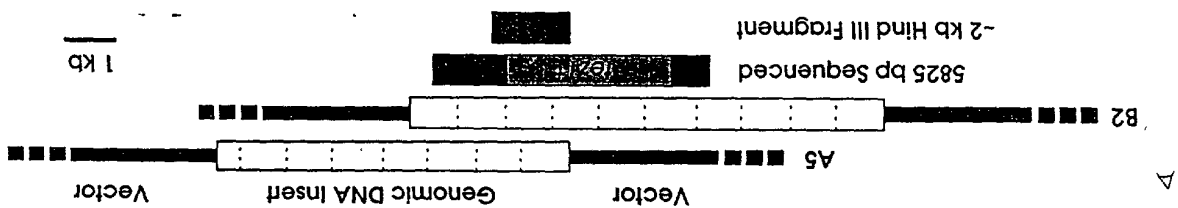


FIGURE 33

FIGURE 34

Poly 4

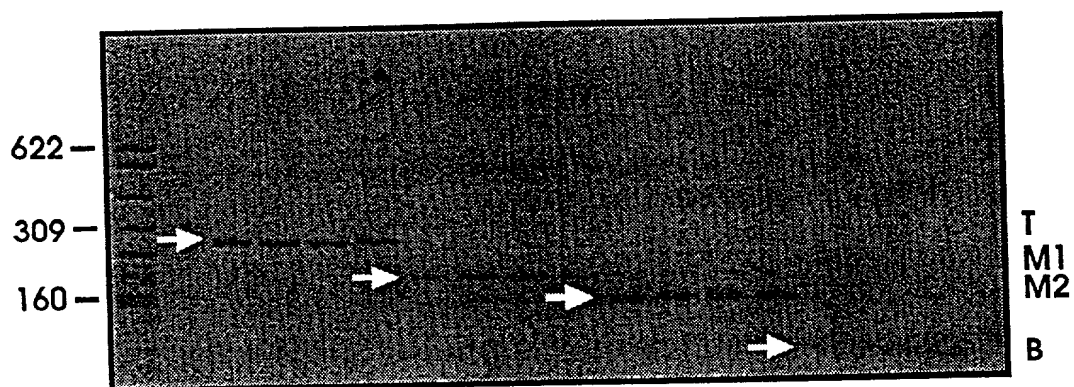
		t		t		c		
	t	a	a	g	c	c	t	c g
5'-	cag	acc	aaa	gga	att	cca	taa	gg -3'
	Q	T	K	G	I	P	Q	G

4 (B')

5 (c')

	D	D	Y	L	L	I	T	
3'-	ctg	ctg	atg	gag	gag	tag	tgg	-5'
	a	a	a	a	a	a	a	
				t	t	t	t	
						c	c	
						<u>Poly 1</u>		

FIGURE 35



Motif B' (4)
QTKGIPQG

Motif C (5)
DDYLLIT

FIGURE 36

PCR Product M2 showed Reasonable Match with Other Telomerase Proteins

```
Ot          LCVSYILSSFYYANLEENALQFLRKESMDPEKPETNLLMRLT
Ea_p123     KGIPQGLCVSSILSSFYYATLESSLGFLRDESMNPENPNVNLLMRLTDDYLLIT
Sp_M2       SILSSFLCHFYMEDLIDEYLSFTKKK-----GSVLLRVV
Sc_p103     DGLFQGSSLSAPIVDLVYDDLLEFYSEFKASPS-----QDTLILKLADDFLIIS
              . * .           * . *           . . . . .
```

```
Q K V G I P Q G
Caa aaa gtt ggt atc cct cag gg..... <---Actual Genomic Sequence.
```

Poly 4

```
      t      t      c
t a a g c c t c g
Cag acc aaa gga att cca taa gg ---->

ag acc aaa gga att cca tca ggC TCA ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG
tc tgg ttt cct taa ggt agt ccG AGT TAA GAC AGT AGA AAA AAC ACA GTA AAG ATA TAC

      K G I P S G S I L S S F L C H F Y M
```

```
GAA GAT TTG ATT GAT GAA TAC CTA TCG TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA
CTT CTA AAC TAA CTA CTT ATG GAT AGC AAA TGC TTT TTC TTT CCT AGT CAC AAC AAT GCT
```

```
E D L I D E Y L S F T K K K G S V L L R
```

```
GTA GTC gac gac tac ctc ctc atc acc
CAT CAG ctg ctg atg gag gag tag tgg
```

```
V V D D Y L L I T
```

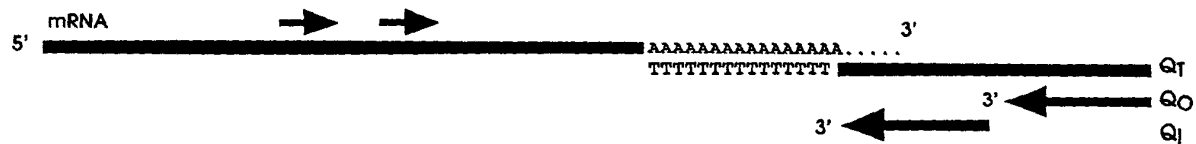
```
<---- ctg ctg atg gag gag tag tgg
      a a a a a a a a
              t t t t
              c c
```

Poly 1

```
.....gac gat ttc ctc ttt ata aca..... <---Actual Genomic Sequence.
      D D F L F I T
```

FIGURE 37

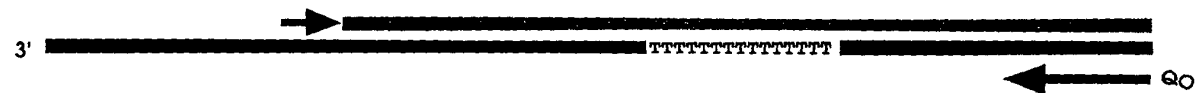
3' RT PCR Strategy



1. Synthesis of cDNA with Q_T Primer.



2. First Round PCR Using Outside Primer and Q_O Primer.



3. Second Round PCR Using Inside Primer and Q_I Primer.



4. Sequence Second Round PCR Products Using Inside Primer or Q_I Primer.



FIGURE 38

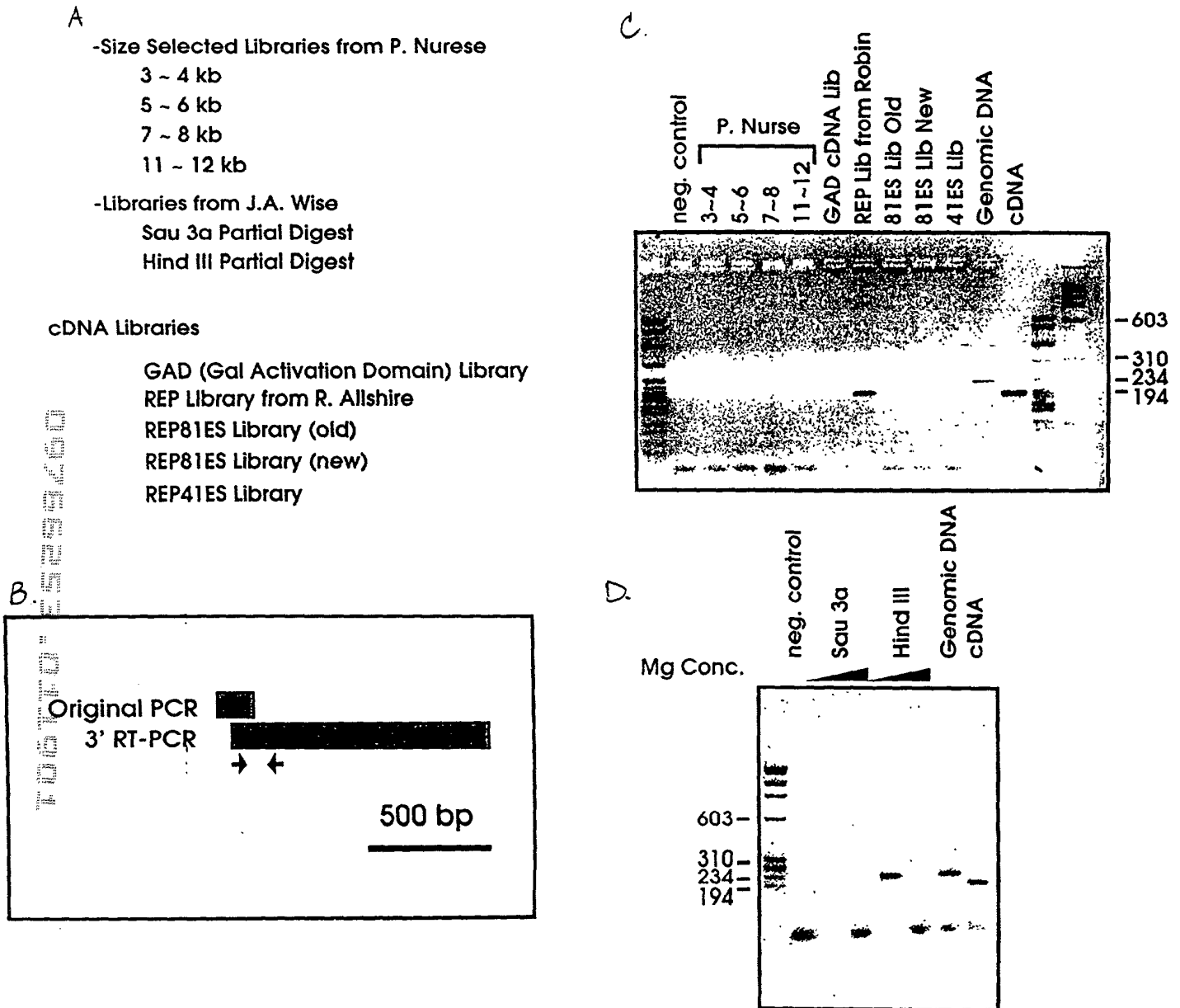


FIGURE 39

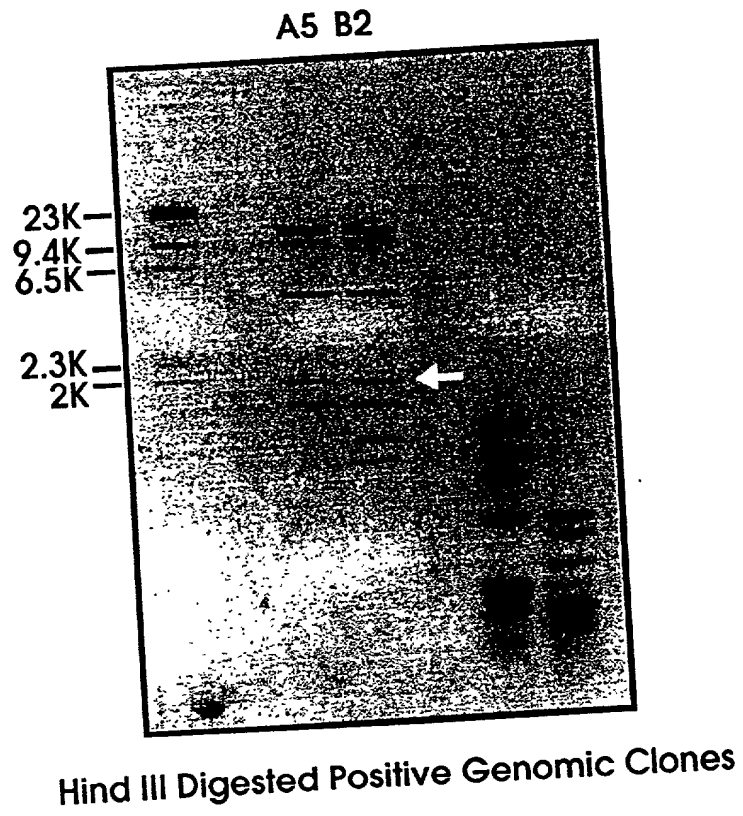
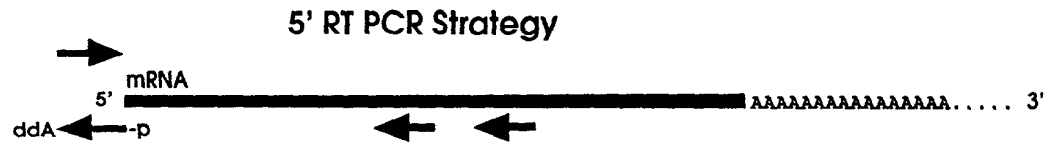


FIGURE 40



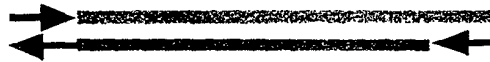
1. Synthesis of cDNA with Specific Downstream Primer.



2. Ligate Oligo with 5'-P and blocked 3' to cDNA using T4 RNA Ligase.



3. First Round PCR



4. Second Round PCR

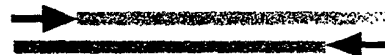


FIGURE 41

Alignment of RT Domains from Telomerase Catalytic Subunits.

```

                                Motif O
S.p. Tez1p (429). WLYNSFIIPILQSFFYITESSDLNRNTVYFRKDIW ... (35) ...
S.c. Est2p (366). WLFRLIPKIIQTFFYCTEISSTVT-IVYFRHDTW ... (35) ...
E.a. p123 (441). WIFEDLVVSLIRCFFYVTEQQKSYSKTYYYRKNIW ... (35) ...
      *          *** **          * * *

      Motif 1      Motif 2      K
      p hh h K      hR h      R
S.p. Tez1p AVIRLLPKK--NTFRLITN-LRKRF ... (61) ...
S.c. Est2p SKMRIIPKKSNEFRIIAIPCRGAD ... (62) ...
E.a. p123 GKLRLIPKK--TTFRPIMTFNKKIV ... (61) ...
      * ***      ** *

      Motif 3(A) AF
      h hDh GY h
S.p. Tez1p KKYFVRIDIKSCYDRIKQDLMFRIVK ... (89) ...
S.c. Est2p ELYFMKFDVKSCYDSIPRMECMRILK ... (75) ...
E.a. p123 KLFFATMDIEKCYDSVNREKLSTFLK ... (107) ...
      * *      ***      *

      Motif 4(B')
      hPQG pP hh h
S.p. Tez1p YLQKVGIPQGSILSSFLCHFYMEDLIDEYLSF ... (6) ...
S.c. Est2p YIREDGLFQGSSLSAPIVDLVYDDLLEFYSEF ... (8) ...
E.a. p123 YKQTKGIPQGLCVSSILSSFYYATLEESSLGF ... (14) ...
      * * ** *      *      *

      Y Motif 5(C)      Motif 6(D)
      h F DDhhh      Gh h cK h
S.p. Tez1p VLLRVVDDFLFITVNKKDAKKFLNLSLRGFEEKHNFSTSLKTVINFENS . (205)
S.c. Est2p LILKLADDFLIISTDQQQVINIKKLAMGGFQKYNANRDKILAVSSQS . (173)
E.a. p123 LLMRLTDDYLLITTQENNAVLFIKLNVSRENGFKFNMKKLQTSFPLS . (209)
      ** * *      *      *

```

FIGURE 42

A

Sp_Tip1p 1 -----MTEHHTPKSRILRFLNQYVYLCT 24
Sc_Est2p 1 -----MTEHHTPKSRILRFLNQYVYLCT 24
Ea_p123 1 MEVDYDQADNHQINSALKTCCEIEKATLYSW 35

Sp_Tip1p 25 LNDYVQLVLRGSPA S SNICERLRSDVQTSFS 37
Sc_Est2p 25 LNDYVQLVLRGSPA S SNICERLRSDVQTSFS 37
Ea_p123 25 LNDYVQLVLRGSPA S SNICERLRSDVQTSFS 37

Sp_Tip1p 38 FFKSTYVQFDPKDFQVQSSPKCQSSAN 40
Sc_Est2p 38 FFKSTYVQFDPKDFQVQSSPKCQSSAN 40
Ea_p123 38 FFKSTYVQFDPKDFQVQSSPKCQSSAN 40

Sp_Tip1p 41 VVKOMFDESERRR NLLMKFSNMHEDFRAMH 122
Sc_Est2p 41 VVKOMFDESERRR NLLMKFSNMHEDFRAMH 122
Ea_p123 41 VVKOMFDESERRR NLLMKFSNMHEDFRAMH 122

Sp_Tip1p 123 VNGVONDVSTFPNYLISILESKNQLLLEIF 126
Sc_Est2p 123 VNGVONDVSTFPNYLISILESKNQLLLEIF 126
Ea_p123 123 VNGVONDVSTFPNYLISILESKNQLLLEIF 126

Sp_Tip1p 129 SDNHYVLSGSEALPNDNYLSIPLKKN 156
Sc_Est2p 129 SDNHYVLSGSEALPNDNYLSIPLKKN 156
Ea_p123 129 SDNHYVLSGSEALPNDNYLSIPLKKN 156

Sp_Tip1p 160 NVFEETYSKKRKRTIETSITN SSKARKE 216
Sc_Est2p 160 NVFEETYSKKRKRTIETSITN SSKARKE 216
Ea_p123 160 NVFEETYSKKRKRTIETSITN SSKARKE 216

Sp_Tip1p 219 WNSISIRFSIFRYSYKQDYVQLHLCID 261
Sc_Est2p 219 WNSISIRFSIFRYSYKQDYVQLHLCID 261
Ea_p123 219 WNSISIRFSIFRYSYKQDYVQLHLCID 261

Sp_Tip1p 252 RNTVHMWLOWIFPROFLINAFQVKQLHKKYIPL 254
Sc_Est2p 252 RNTVHMWLOWIFPROFLINAFQVKQLHKKYIPL 254
Ea_p123 252 RNTVHMWLOWIFPROFLINAFQVKQLHKKYIPL 254

Sp_Tip1p 266 VS-----OSTYVPRRLVPLLEQTAQLH 315
Sc_Est2p 266 VS-----OSTYVPRRLVPLLEQTAQLH 315
Ea_p123 266 VS-----OSTYVPRRLVPLLEQTAQLH 315

Sp_Tip1p 343 LSKVYHYVYIDYHDEKILSYSLKPNQ 342
Sc_Est2p 343 LSKVYHYVYIDYHDEKILSYSLKPNQ 342
Ea_p123 343 LSKVYHYVYIDYHDEKILSYSLKPNQ 342

Sp_Tip1p 343 SKYYEELFSYTTDNKCTOHEFFYNI 374
Sc_Est2p 343 SKYYEELFSYTTDNKCTOHEFFYNI 374
Ea_p123 343 SKYYEELFSYTTDNKCTOHEFFYNI 374

Sp_Tip1p 360 WONGRIFEILILETFLKSRYESFSLHYMS 362
Sc_Est2p 360 WONGRIFEILILETFLKSRYESFSLHYMS 362
Ea_p123 360 WONGRIFEILILETFLKSRYESFSLHYMS 362

Sp_Tip1p 363 NIKISEIEVLQKRSNAKMLCSKRRKIFA 405
Sc_Est2p 363 NIKISEIEVLQKRSNAKMLCSKRRKIFA 405
Ea_p123 363 NIKISEIEVLQKRSNAKMLCSKRRKIFA 405

Sp_Tip1p 426 EFIVLYNSFIPILOS SSSDLNRRTV 456
Sc_Est2p 426 EFIVLYNSFIPILOS SSSDLNRRTV 456
Ea_p123 426 EFIVLYNSFIPILOS SSSDLNRRTV 456

Sp_Tip1p 458 FKDIKLLCRPFSYMKMEAFKINENVRMD 460
Sc_Est2p 458 FKDIKLLCRPFSYMKMEAFKINENVRMD 460
Ea_p123 458 FKDIKLLCRPFSYMKMEAFKINENVRMD 460

Sp_Tip1p 462 YOKTTLPAVILLLSNHNLTLNRKRF 522
Sc_Est2p 462 YOKTTLPAVILLLSNHNLTLNRKRF 522
Ea_p123 462 YOKTTLPAVILLLSNHNLTLNRKRF 522

Sp_Tip1p 523 IKMQSHKKMLVSTNOTLRPYASIKHIL 524
Sc_Est2p 523 IKMQSHKKMLVSTNOTLRPYASIKHIL 524
Ea_p123 523 IKMQSHKKMLVSTNOTLRPYASIKHIL 524

Sp_Tip1p 553 ESSQIPFNLEYVMKLLTFKKDLHRRNFR 564
Sc_Est2p 553 ESSQIPFNLEYVMKLLTFKKDLHRRNFR 564
Ea_p123 553 ESSQIPFNLEYVMKLLTFKKDLHRRNFR 564

Sp_Tip1p 566 VVRIILKSRIKQDLMFRIYKLLKOP 616
Sc_Est2p 566 VVRIILKSRIKQDLMFRIYKLLKOP 616
Ea_p123 566 VVRIILKSRIKQDLMFRIYKLLKOP 616

Sp_Tip1p 617 VVKYATIHATSBRATKN 624
Sc_Est2p 617 VVKYATIHATSBRATKN 624
Ea_p123 617 VVKYATIHATSBRATKN 624

Sp_Tip1p 636 VSEAFSYFDMVPEKVLLOS-MKTSDF 644
Sc_Est2p 636 VSEAFSYFDMVPEKVLLOS-MKTSDF 644
Ea_p123 636 VSEAFSYFDMVPEKVLLOS-MKTSDF 644

Sp_Tip1p 666 FFDYWTSSSEIFKMLKHLNLSGHIVKIGNSO 668
Sc_Est2p 666 FFDYWTSSSEIFKMLKHLNLSGHIVKIGNSO 668
Ea_p123 666 FFDYWTSSSEIFKMLKHLNLSGHIVKIGNSO 668

Sp_Tip1p 688 LOKVIFPILSPLCNFYMEDIDEVLS 721
Sc_Est2p 688 LOKVIFPILSPLCNFYMEDIDEVLS 721
Ea_p123 688 LOKVIFPILSPLCNFYMEDIDEVLS 721

Sp_Tip1p 722 KKG-----SVLLRVYFIVYKDKDAKK 756
Sc_Est2p 722 KKG-----SVLLRVYFIVYKDKDAKK 756
Ea_p123 722 KKG-----SVLLRVYFIVYKDKDAKK 756

Sp_Tip1p 757 FLNLSLRGFEKHNFTSLTYVINFEK 766
Sc_Est2p 757 FLNLSLRGFEKHNFTSLTYVINFEK 766
Ea_p123 757 FLNLSLRGFEKHNFTSLTYVINFEK 766

Sp_Tip1p 767 -----LHNTTFESKKRPFVFNHRS 816
Sc_Est2p 767 -----LHNTTFESKKRPFVFNHRS 816
Ea_p123 767 -----LHNTTFESKKRPFVFNHRS 816

Sp_Tip1p 850 AFAQFIDFNSFNCHYIRGYNCMR 862
Sc_Est2p 850 AFAQFIDFNSFNCHYIRGYNCMR 862
Ea_p123 850 AFAQFIDFNSFNCHYIRGYNCMR 862

Sp_Tip1p 864 AQAFLRMILPORMFIDLLNVIQRKWK 916
Sc_Est2p 864 AQAFLRMILPORMFIDLLNVIQRKWK 916
Ea_p123 864 AQAFLRMILPORMFIDLLNVIQRKWK 916

Sp_Tip1p 916 LAEILQYTSRRFLSSAEVYFLGLQMR 946
Sc_Est2p 916 LAEILQYTSRRFLSSAEVYFLGLQMR 946
Ea_p123 916 LAEILQYTSRRFLSSAEVYFLGLQMR 946

Sp_Tip1p 949 FKYPHCFEQLIYQSLTDILKPLRPV 961
Sc_Est2p 949 FKYPHCFEQLIYQSLTDILKPLRPV 961
Ea_p123 949 FKYPHCFEQLIYQSLTDILKPLRPV 961

Sp_Tip1p 962 LHRRIAD 964
Sc_Est2p 962 LHRRIAD 964
Ea_p123 962 LHRRIAD 964

B

Sp_Tip1p 1 -----MTEHHTPKSRILRFLNQYVYLCT 24
Sc_Est2p 1 -----MTEHHTPKSRILRFLNQYVYLCT 24
Ea_p123 1 MEVDYDQADNHQINSALKTCCEIEKATLYSW 35

Sp_Tip1p 26 LNDYVQLVLRGSPA S SNICERLRSDVQTSFS 37
Sc_Est2p 26 LNDYVQLVLRGSPA S SNICERLRSDVQTSFS 37
Ea_p123 26 LNDYVQLVLRGSPA S SNICERLRSDVQTSFS 37

Sp_Tip1p 38 FFKSTYVQFDPKDFQVQSSPKCQSSAN 40
Sc_Est2p 38 FFKSTYVQFDPKDFQVQSSPKCQSSAN 40
Ea_p123 38 FFKSTYVQFDPKDFQVQSSPKCQSSAN 40

Sp_Tip1p 41 VVKOMFDESERRR NLLMKFSNMHEDFRAMH 122
Sc_Est2p 41 VVKOMFDESERRR NLLMKFSNMHEDFRAMH 122
Ea_p123 41 VVKOMFDESERRR NLLMKFSNMHEDFRAMH 122

Sp_Tip1p 123 VNGVONDVSTFPNYLISILESKNQLLLEIF 126
Sc_Est2p 123 VNGVONDVSTFPNYLISILESKNQLLLEIF 126
Ea_p123 123 VNGVONDVSTFPNYLISILESKNQLLLEIF 126

Sp_Tip1p 129 SDNHYVLSGSEALPNDNYLSIPLKKN 156
Sc_Est2p 129 SDNHYVLSGSEALPNDNYLSIPLKKN 156
Ea_p123 129 SDNHYVLSGSEALPNDNYLSIPLKKN 156

Sp_Tip1p 160 NVFEETYSKKRKRTIETSITN SSKARKE 216
Sc_Est2p 160 NVFEETYSKKRKRTIETSITN SSKARKE 216
Ea_p123 160 NVFEETYSKKRKRTIETSITN SSKARKE 216

Sp_Tip1p 219 WNSISIRFSIFRYSYKQDYVQLHLCID 261
Sc_Est2p 219 WNSISIRFSIFRYSYKQDYVQLHLCID 261
Ea_p123 219 WNSISIRFSIFRYSYKQDYVQLHLCID 261

Sp_Tip1p 252 RNTVHMWLOWIFPROFLINAFQVKQLHKKYIPL 254
Sc_Est2p 252 RNTVHMWLOWIFPROFLINAFQVKQLHKKYIPL 254
Ea_p123 252 RNTVHMWLOWIFPROFLINAFQVKQLHKKYIPL 254

Sp_Tip1p 266 VS-----OSTYVPRRLVPLLEQTAQLH 315
Sc_Est2p 266 VS-----OSTYVPRRLVPLLEQTAQLH 315
Ea_p123 266 VS-----OSTYVPRRLVPLLEQTAQLH 315

Sp_Tip1p 343 LSKVYHYVYIDYHDEKILSYSLKPNQ 342
Sc_Est2p 343 LSKVYHYVYIDYHDEKILSYSLKPNQ 342
Ea_p123 343 LSKVYHYVYIDYHDEKILSYSLKPNQ 342

Sp_Tip1p 343 SKYYEELFSYTTDNKCTOHEFFYNI 374
Sc_Est2p 343 SKYYEELFSYTTDNKCTOHEFFYNI 374
Ea_p123 343 SKYYEELFSYTTDNKCTOHEFFYNI 374

Sp_Tip1p 360 WONGRIFEILILETFLKSRYESFSLHYMS 362
Sc_Est2p 360 WONGRIFEILILETFLKSRYESFSLHYMS 362
Ea_p123 360 WONGRIFEILILETFLKSRYESFSLHYMS 362

Sp_Tip1p 363 NIKISEIEVLQKRSNAKMLCSKRRKIFA 405
Sc_Est2p 363 NIKISEIEVLQKRSNAKMLCSKRRKIFA 405
Ea_p123 363 NIKISEIEVLQKRSNAKMLCSKRRKIFA 405

Sp_Tip1p 426 EFIVLYNSFIPILOS SSSDLNRRTV 456
Sc_Est2p 426 EFIVLYNSFIPILOS SSSDLNRRTV 456
Ea_p123 426 EFIVLYNSFIPILOS SSSDLNRRTV 456

Sp_Tip1p 458 FKDIKLLCRPFSYMKMEAFKINENVRMD 460
Sc_Est2p 458 FKDIKLLCRPFSYMKMEAFKINENVRMD 460
Ea_p123 458 FKDIKLLCRPFSYMKMEAFKINENVRMD 460

Sp_Tip1p 462 YOKTTLPAVILLLSNHNLTLNRKRF 522
Sc_Est2p 462 YOKTTLPAVILLLSNHNLTLNRKRF 522
Ea_p123 462 YOKTTLPAVILLLSNHNLTLNRKRF 522

Sp_Tip1p 523 IKMQSHKKMLVSTNOTLRPYASIKHIL 524
Sc_Est2p 523 IKMQSHKKMLVSTNOTLRPYASIKHIL 524
Ea_p123 523 IKMQSHKKMLVSTNOTLRPYASIKHIL 524

Sp_Tip1p 553 ESSQIPFNLEYVMKLLTFKKDLHRRNFR 564
Sc_Est2p 553 ESSQIPFNLEYVMKLLTFKKDLHRRNFR 564
Ea_p123 553 ESSQIPFNLEYVMKLLTFKKDLHRRNFR 564

Sp_Tip1p 566 VVRIILKSRIKQDLMFRIYKLLKOP 616
Sc_Est2p 566 VVRIILKSRIKQDLMFRIYKLLKOP 616
Ea_p123 566 VVRIILKSRIKQDLMFRIYKLLKOP 616

Sp_Tip1p 617 VVKYATIHATSBRATKN 624
Sc_Est2p 617 VVKYATIHATSBRATKN 624
Ea_p123 617 VVKYATIHATSBRATKN 624

Sp_Tip1p 636 VSEAFSYFDMVPEKVLLOS-MKTSDF 644
Sc_Est2p 636 VSEAFSYFDMVPEKVLLOS-MKTSDF 644
Ea_p123 636 VSEAFSYFDMVPEKVLLOS-MKTSDF 644

Sp_Tip1p 666 FFDYWTSSSEIFKMLKHLNLSGHIVKIGNSO 668
Sc_Est2p 666 FFDYWTSSSEIFKMLKHLNLSGHIVKIGNSO 668
Ea_p123 666 FFDYWTSSSEIFKMLKHLNLSGHIVKIGNSO 668

Sp_Tip1p 688 LOKVIFPILSPLCNFYMEDIDEVLS 721
Sc_Est2p 688 LOKVIFPILSPLCNFYMEDIDEVLS 721
Ea_p123 688 LOKVIFPILSPLCNFYMEDIDEVLS 721

Sp_Tip1p 722 KKG-----SVLLRVYFIVYKDKDAKK 756
Sc_Est2p 722 KKG-----SVLLRVYFIVYKDKDAKK 756
Ea_p123 722 KKG-----SVLLRVYFIVYKDKDAKK 756

Sp_Tip1p 757 FLNLSLRGFEKHNFTSLTYVINFEK 766
Sc_Est2p 757 FLNLSLRGFEKHNFTSLTYVINFEK 766
Ea_p123 757 FLNLSLRGFEKHNFTSLTYVINFEK 766

Sp_Tip1p 767 -----LHNTTFESKKRPFVFNHRS 816
Sc_Est2p 767 -----LHNTTFESKKRPFVFNHRS 816
Ea_p123 767 -----LHNTTFESKKRPFVFNHRS 816

Sp_Tip1p 850 AFAQFIDFNSFNCHYIRGYNCMR 862
Sc_Est2p 850 AFAQFIDFNSFNCHYIRGYNCMR 862
Ea_p123 850 AFAQFIDFNSFNCHYIRGYNCMR 862

Sp_Tip1p 864 AQAFLRMILPORMFIDLLNVIQRKWK 916
Sc_Est2p 864 AQAFLRMILPORMFIDLLNVIQRKWK 916
Ea_p123 864 AQAFLRMILPORMFIDLLNVIQRKWK 916

Sp_Tip1p 916 LAEILQYTSRRFLSSAEVYFLGLQMR 946
Sc_Est2p 916 LAEILQYTSRRFLSSAEVYFLGLQMR 946
Ea_p123 916 LAEILQYTSRRFLSSAEVYFLGLQMR 946

Sp_Tip1p 949 FKYPHCFEQLIYQSLTDILKPLRPV 961
Sc_Est2p 949 FKYPHCFEQLIYQSLTDILKPLRPV 961
Ea_p123 949 FKYPHCFEQLIYQSLTDILKPLRPV 961

Sp_Tip1p 962 LHRRIAD 964
Sc_Est2p 962 LHRRIAD 964
Ea_p123 962 LHRRIAD 964

FIGURE 43

Disruption strategy for the putative telomerase genes.

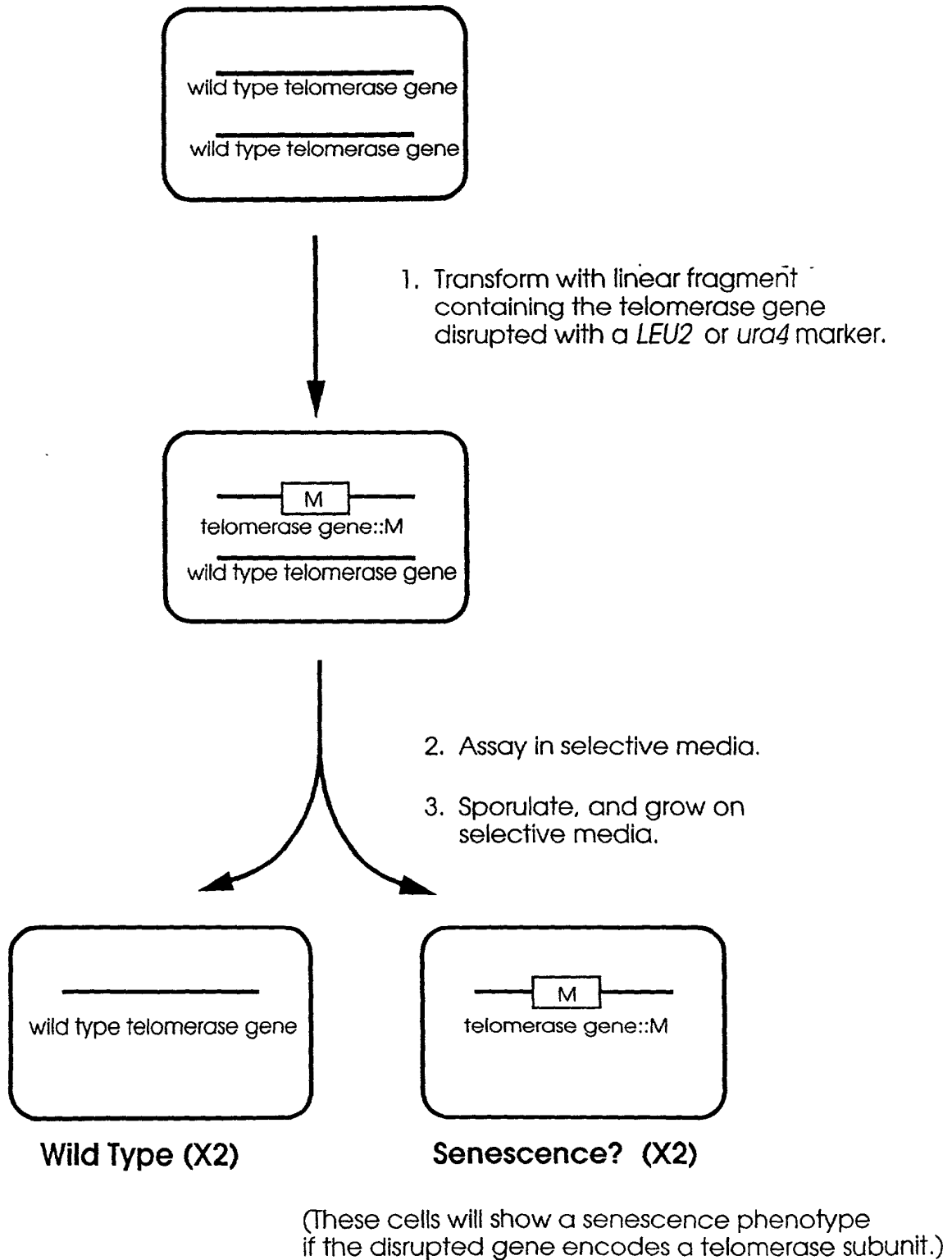


FIGURE 44

An Example of Confirmation of *tez1* disruption By PCR

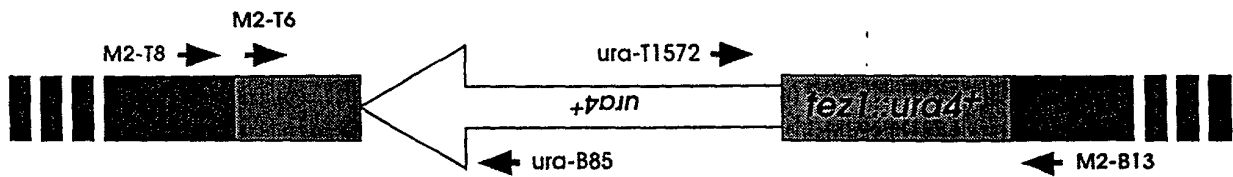
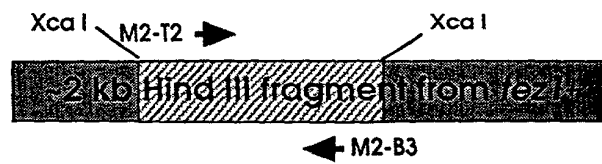
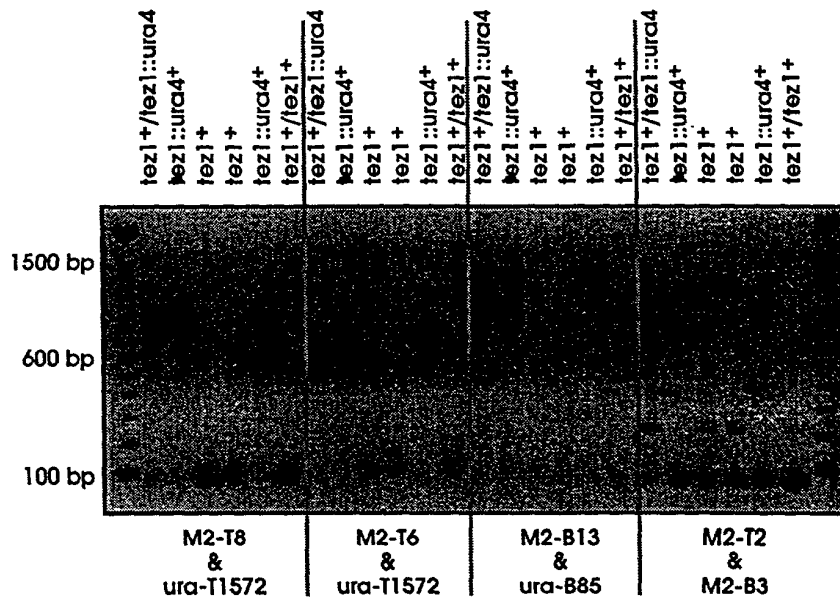


FIGURE 45

Tez1 disruption causes progressive shortening of telomeres in *S. pombe*

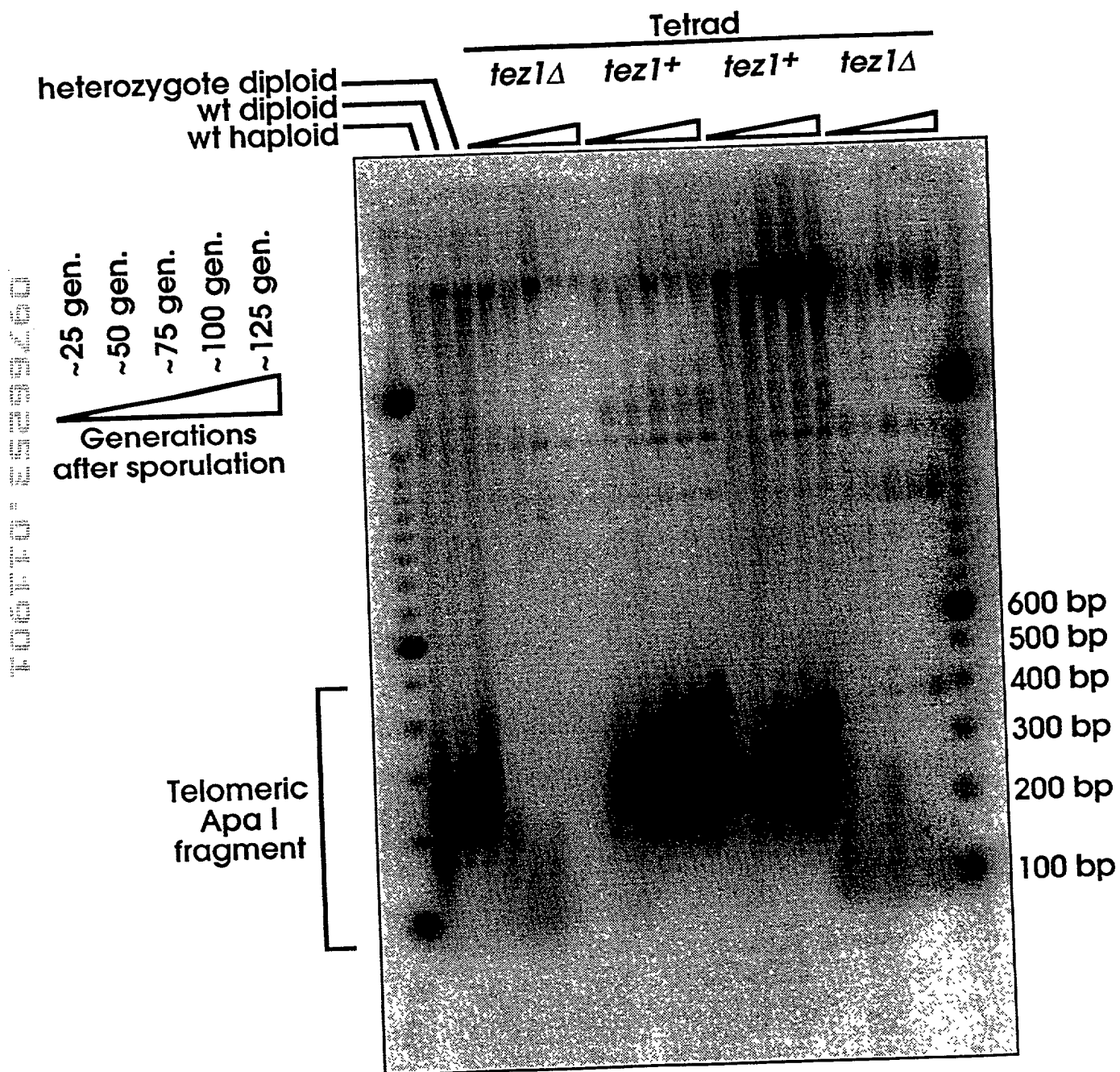


FIGURE 46

1 ggtaccgatttactttcctttcttcataagctaattgcttcctcgaacgctcctaaatctctggaaatatttttacaaga 80
81 actcaataacaataccaagtcaaattccaatatgaagggtgttattagtgatcgataatatttctattttatcggtcgtta 160
161 ccaagtataaggacaaaaagaacaacttccttccccctaaagacttttactttattaattttacttttcaaataatatttcg 240
241 ggttcgcttacttttaatecgtggtaactgttttagctgctacttctagccaaccgcgtgtttctaccccgctcattggatat 320
321 agctcttggagtagctcacagaaatccttacaaatcttctgatgagactatattagattcattacagtcctgcatattc 400
401 ttaacatggagccttacacttttagatgagtcacgtcgcgatgatggagtatttggtatcatccaacgtttgcccttgaaaag 480
481 gttgataattatttgcaaaatcatgtccttagtggtggtaatccgcgaaagtttttgatgcttgacacgtctagcatg 560
561 attgagatattcaaaaatttctatccactacaactcctttaacgcgggttttatttttctattttctattctcatgttggt 640
641 ccaaataatgtatcatctcgtattaggctttttccggttttactcctggaatcgtaccttttccactattccccctaata 720
721 ataactctaaattagtttcgcttataattgatagtagtagaaagattgggtgattctactcgtgtaattgttattagtttaa 800
801 gatactttgcaaaacatttattagctatcattatataaaaaaaatcctataattataaatattaatcaatatttgcgggc 880
881 actatttattttaaacggttatgatcagtaggacactttgcatatatatagttatgcttaatgggttacttgtaacttgc 958

959 ATG ACC GAA CAC CAT ACC CCC AAA AGC AGG ATT CTT CGC TTT CTA GAG AAT CAA TAT GTA 1018
1 M T E H H T P K S R I L R F L E N Q Y V 20

1019 TAC CTA TGT ACC TTA AAT GAT TAT GTA CAA CTT GTT TTG AGA GGG TCG CCG GCA AGC TCG 1078
21 Y L C T L N D Y V Q L V L R G S P A S S 40

1079 TAT AGC AAT ATA TGC GAA CGC TTG AGA AGC GAT GTA CAA ACG TCC TTT TCT ATT TTT CTT 1138
41 Y S N I C E R L R S D V Q T S F S I F L 60

1139 CAT TCG ACT GTA GTC GGC TTC GAC AGT AAG CCA GAT GAA GGT GTT CAA TTT TCT TCT CCA 1198
61 H S T V V G F D S K P D E G V Q F S S P 80

FIGURE 46 (cont.)

1199	AAA TGC TCA CAG TCA GAG	gtatatatatttttgttttgatttttttctattcgggatagctaatatatgggcag	1272
81	K C S Q S E		86
1273	CTA ATA GCG AAT GTT GTA AAA CAG ATG TTC GAT GAA AGT TTT GAG CGT CGA AGG AAT CTA		1332
87	L I A N V V K Q M F D E S F E R R R N L		106
1333	CTG ATG AAA GGG TTT TCC ATG	gtaaggatttctaattgtgaaatattttacctgcaattactgtttcaaagaga	1405
107	L M K G F S M		113
1406	ttgtatttaaccgataaag	AAT CAT GAA GAT TTT CGA GCC ATG CAT GTA AAC GGA GTA CAA AAT	1469
114		N H E D F R A M H V N G V Q N	128
1470	GAT CTC GTT TCT ACT TTT CCT AAT TAC CTT ATA TCT ATA CTT GAG TCA AAA AAT TGG CAA		1529
129	D L V S T F P N Y L I S I L E S K N W Q		148
1530	CTT TTG TTA GAA AT	gtaaataaccggttaagatgttgcgcactttgaacaagactgacaagtatag T ATC GGC	1601
149	L L L E I		I G 155
1602	AGT GAT GCC ATG CAT TAC TTA TTA TCC AAA GGA AGT ATT TTT GAG GCT CTT CCA AAT GAC		1661
156	S D A M H Y L L S K G S I F E A L P N D		175
1662	AAT TAC CTT CAG ATT TCT GGC ATA CCA CTT TTT AAA AAT AAT GTG TTT GAG GAA ACT GTG		1721
176	N Y L Q I S G I P L F K N N V F E E T V		195
1722	TCA AAA AAA AGA AAG CGA ACC ATT GAA ACA TCC ATT ACT CAA AAT AAA AGC GCC CGC AAA		1781
196	S K K R K R T I E T S I T Q N K S A R K		215
1782	GAA GTT TCC TGG AAT AGC ATT TCA ATT AGT AGG TTT AGC ATT TTT TAC AGG TCA TCC TAT		1841
216	E V S W N S I S I S R F S I F Y R S S Y		235
1842	AAG AAG TTT AAG CAA G	gtaactaatactgttatccttcataactaatttttag AT CTA TAT TTT AAC	
1907			
236	K K F K Q D		L Y F N 245
1908	TTA CAC TCT ATT TGT GAT CGG AAC ACA GTA CAC ATG TGG CTT CAA TGG ATT TTT CCA AGG		1967
246	L H S I C D R N T V H M W L Q W I F P R		265
1968	CAA TTT GGA CTT ATA AAC GCA TTT CAA GTG AAG CAA TTG CAC AAA GTG ATT CCA CTG GTA		2027
266	Q F G L I N A F Q V K Q L H K V I P L V		285
2028	TCA CAG AGT ACA GTT GTG CCC AAA CGT CTC CTA AAG GTA TAC CCT TTA ATT GAA CAA ACA		2087
286	S Q S T V V P K R L L K V Y P L I E Q T		305
2088	GCA AAG CGA CTC CAT CGT ATT TCT CTA TCA AAA GTT TAC AAC CAT TAT TGC CCA TAT ATT		2147
306	A K R L H R I S L S K V Y N H Y C P Y I		325
2148	GAC ACC CAC GAT GAT GAA AAA ATC CTT AGT TAT TCC TTA AAG CCG AAC CAG GTG TTT GCG		2207
326	D T H D D E K I L S Y S L K P N Q V F A		345
2208	TTT CTT CGA TCC ATT CTT GTT CGA GTG TTT CCT AAA TTA ATC TGG GGT AAC CAA AGG ATA		2267
346	F L R S I L V R V F P K L I W G N Q R I		365
2268	TTT GAG ATA ATA TTA AAA G	gtattgtataaaatttattaccactaaccgattttaccag AC CTC GAA ACT	2336
366	F E I I L K D		L E T 375

FIGURE 46 (cont.)

2337	TTC	TTG	AAA	TTA	TCG	AGA	TAC	GAG	TCT	TTT	AGT	TTA	CAT	TAT	TTA	ATG	AGT	AAC	ATA	AAG	2396																	
376	F	L	K	L	S	R	Y	E	S	F	S	L	H	Y	L	M	S	N	I	K	395																	
2397	gtaatatgccaaatttttttaccattaattaacaatcag																			ATT	TCA	GAA	ATT	GAA	TGG	CTA	GTC	CTT	GGA	2465								
396																				I	S	E	I	E	W	L	V	L	G	405								
2466	AAA	AGG	TCA	AAT	GCG	AAA	ATG	TGC	TTA	AGT	GAT	TTT	GAG	AAA	CGC	AAG	CAA	ATA	TTT	GCG	2525																	
406	K	R	S	N	A	K	M	C	L	S	D	F	E	K	R	K	Q	I	F	A	425																	
2526	GAA	TTC	ATC	TAC	TGG	CTA	TAC	AAT	TCG	TTT	ATA	ATA	CCT	ATT	TTA	CAA	TCT	TTT	TTT	TAT	2585																	
426	E	F	I	Y	W	L	Y	N	S	F	I	I	P	I	L	Q	S	F	F	Y	445																	
2586	ATC	ACT	GAA	TCA	AGT	GAT	TTA	CGA	AAT	CGA	ACT	GTT	TAT	TTT	AGA	AAA	GAT	ATT	TGG	AAA	2645																	
446	I	T	E	S	S	D	L	R	N	R	T	V	Y	F	R	K	D	I	W	K	465																	
2646	CTC	TTG	TGC	CGA	CCC	TTT	ATT	ACA	TCA	ATG	AAA	ATG	GAA	GCG	TTT	GAA	AAA-ATA	AAC	GAG	2705																		
466	L	L	C	R	P	F	I	T	S	M	K	M	E	A	F	E	K	I	N	E	485																	
2706	gtatttttaaagtattttttgcaaaaagctaataatttttcag																			AAC	AAT	GTT	AGG	ATG	GAT	ACT	CAG	AAA	ACT	2775								
486																				N	N	V	R	M	D	T	Q	K	T	495								
2776	ACT	TTG	CCT	CCA	GCA	GTT	ATT	CGT	CTA	TTA	CCT	AAG	AAG	AAT	ACC	TTT	CGT	CTC	ATT	ACG	2835																	
496	T	L	P	P	A	V	I	R	L	L	P	K	K	N	T	F	R	L	I	T	515																	
2836	AAT	TTA	AGA	AAA	AGA	TTC	TTA	ATA	AAG	gtattaatttttggatcatcaatgtactttacttctaatactatta											2906																	
516	N	L	R	K	R	F	L	I	K												524																	
2907	ttagcag																			ATG	GGT	TCA	AAC	AAA	AAA	ATG	TTA	GTC	AGT	ACG	AAC	CAA	ACT	TTA	CGA	CCT	GTG	2967
525																				M	G	S	N	K	K	M	L	V	S	T	N	Q	T	L	R	P	V	542
2968	GCA	TCG	ATA	CTG	AAA	CAT	TTA	ATC	AAT	GAA	GAA	AGT	AGT	GGT	ATT	CCA	TTT	AAC	TTG	GAG	3027																	
543	A	S	I	L	K	H	L	I	N	E	E	S	S	G	I	P	F	N	L	E	562																	
3028	GTT	TAC	ATG	AAG	CTT	CTT	ACT	TTT	AAG	AAG	GAT	CTT	CTT	AAG	CAC	CGA	ATG	TTT	GG	gtaat	3088																	
563	V	Y	M	K	L	L	T	F	K	K	D	L	L	K	H	R	M	F	G		581																	
3089	tatataatgcgcgatttcctcattattaatttttgcag																			G	CGT	AAG	AAG	TAT	TTT	GTA	CGG	ATA	GAT	ATA	3155							
582																				R	K	K	Y	F	V	R	I	D	I		591							
3156	AAA	TCC	TGT	TAT	GAT	CGA	ATA	AAG	CAA	GAT	TTG	ATG	TTT	CGG	ATT	GTT	AAA	AAG	AAA	CTC	3215																	
592	K	S	C	Y	D	R	I	K	Q	D	L	M	F	R	I	V	K	K	K	L	611																	
3216	AAG	GAT	CCC	GAA	TTT	GTA	ATT	CGA	AAG	TAT	GCA	ACC	ATA	CAT	GCA	ACA	AGT	GAC	CGA	GCT	3275																	
612	K	D	P	E	F	V	I	R	K	Y	A	T	I	H	A	T	S	D	R	A	631																	
3276	ACA	AAA	AAC	TTT	GTT	AGT	GAG	GCG	TTT	TCC	TAT	T	gtaagtttatttttttcattggaattttttaacaa											3343														
632	T	K	N	F	V	S	E	A	F	S	Y	F												643														
3344	attccttttttag																			TT	GAT	ATG	GTG	CCT	TTT	GAA	AAA	GTC	GTG	CAG	TTA	CTT	TCT	ATG	AAA	ACA	3405	
644																				D	M	V	P	F	E	K	V	V	Q	L	L	S	M	K	T	659		
3406	TCA	GAT	ACT	TTG	TTT	GTT	GAT	TTT	GTG	GAT	TAT	TGG	ACC	AAA	AGT	TCT	TCT	GAA	ATT	TTT	3465																	
660	S	D	T	L	F	V	D	F	V	D	Y	W	T	K	S	S	S	E	I	F	679																	
3466	AAA	ATG	CTC	AAG	GAA	CAT	CTC	TCT	GGA	CAC	ATT	GTT	AAG	gtataccaattgttgaattgtaataaca											3532													
680	K	M	L	K	E	H	L	S	G	H	I	V	K												692													

1

390
79
396
81
402
83
409
84
415
86
421

FIGURE 46 (cont.)

4666 gctgacccccaaagcaagcatactataggatttctagtaaagtaaaattaatctcgttattagttttgattgacttgtct 4745
4746 ttatccttatactttttaagaaagattgacagtgggttgctgactactgccacatgccattaaacgggagtggttaaaca 4825
4826 ttaaaagtaatacatgaggctaatactcctttcatttagaataaggaaagtggttttctataatgaataatgcccgacta 4905
4906 atgcaaaaagacgaagattatcttctaacaagggggattaagcatatccgaaggaaaagagagtaatataccagtggt 4985
4986 gttgaagaaagcaaggataaatttgaacaagcttctgcagatgacaggctaattttggtgaccgaattttggtaaaagc 5065
5066 cccaggttatccatggtggccggccttgctactgagacgaaaagaaactaaggatagtttgaataactaatagctcattta 5145
5146 atgtcttatataagggttttgttttttctgacttcaattttgcatgggtgaaaagaaatagtgttaagccattattggat 5225
5226 tccgaaatagccaaatttcttggttcctcaaagcgggaagtctaaagaacttattgaagcttatgaggcttcaaaaactcc 5305
5306 tcctgatttaaaggaggaatcttccaccgatgaggaaatggatagcttatcagctgctgaggagaagcctaattttttgc 5385
5386 aaaaagaaaatatcattgggagacatctcttgatgaatcagatgctggagagtatctccagcggatccttgatgtcaata 5465
5466 acttctatttctgaaatgtatggtcctactgtcgcttcgacttctcgtagctctacgcagttaagtgaacaaagggtacc 5544

4666 gctgacccccaaagcaagcatactataggatttctagtaaagtaaaattaatctcgttattagttttgattgacttgtct 4745
4746 ttatccttatactttttaagaaagattgacagtgggttgctgactactgccacatgccattaaacgggagtggttaaaca 4825
4826 ttaaaagtaatacatgaggctaatactcctttcatttagaataaggaaagtggttttctataatgaataatgcccgacta 4905
4906 atgcaaaaagacgaagattatcttctaacaagggggattaagcatatccgaaggaaaagagagtaatataccagtggt 4985
4986 gttgaagaaagcaaggataaatttgaacaagcttctgcagatgacaggctaattttggtgaccgaattttggtaaaagc 5065
5066 cccaggttatccatggtggccggccttgctactgagacgaaaagaaactaaggatagtttgaataactaatagctcattta 5145
5146 atgtcttatataagggttttgttttttctgacttcaattttgcatgggtgaaaagaaatagtgttaagccattattggat 5225
5226 tccgaaatagccaaatttcttggttcctcaaagcgggaagtctaaagaacttattgaagcttatgaggcttcaaaaactcc 5305
5306 tcctgatttaaaggaggaatcttccaccgatgaggaaatggatagcttatcagctgctgaggagaagcctaattttttgc 5385
5386 aaaaagaaaatatcattgggagacatctcttgatgaatcagatgctggagagtatctccagcggatccttgatgtcaata 5465
5466 acttctatttctgaaatgtatggtcctactgtcgcttcgacttctcgtagctctacgcagttaagtgaacaaagggtacc 5544

FIGURE 47 (cont.)

190	thr	tyr	cys	val	arg	arg	tyr	ala	val	val	200	gln	lys	ala	ala	met
	ACG	TAC	TGC	GTG	CGT	CGG	TAT	GCC	GTG	GTC		CAG	AAG	GCC	GCC	ATG
	gly	thr	ser	ala	arg	pro	ser	arg	ala	thr	210	ser	tyr	val	gln	cys
	GGC	ACG	TCC	GCA	AGG	CCT	TCA	AGA	GCC	ACG		TCC	TAC	GTC	CAG	TGC
220	gln	gly	ile	pro	gln	gly	ser	ile	leu	ser	230	thr	leu	leu	cys	ser
	CAG	GGG	ATC	CCG	CAG	GGC	TCC	ATC	CTC	TCC		ACG	CTG	CTC	TGC	AGC
	leu	cys	tyr	gly	asp	met	glu	asn	lys	leu	240	phe	ala	gly	ile	arg
	CTG	TGC	TAC	GGC	GAC	ATG	GAG	AAC	AAG	CTG		TTT	GCG	GGG	ATT	CGG
250	arg	asp	gly	leu	leu	leu	arg	leu	val	asp	260	asp	phe	leu	leu	val
	CGG	GAC	GGG	CTG	CTC	CTG	CGT	TTG	GTG	GAT		GAT	TTC	TTG	TTG	GTG
	thr	pro	his	leu	thr	his	ala	lys	thr	phe	270	leu	arg	thr	leu	val
	ACA	CCT	CAC	CTC	ACC	CAC	GCG	AAA	ACC	TTC		CTC	AGG	ACC	CTG	GTC
280	arg	gly	val	pro	glu	tyr	gly	cys	val	val	290	asn	leu	arg	lys	thr
	CGA	GGT	GTC	CCT	GAG	TAT	GGC	TGC	GTG	GTG		AAC	TTG	CGG	AAG	ACA
	val	val	asn	phe	pro	val	glu	asp	glu	ala	300	leu	gly	gly	thr	ala
	GTG	GTG	AAC	TTC	CCT	GTA	GAA	GAC	GAG	GCC		CTG	GGT	GGC	ACG	GCT
310	phe	val	gln	met	pro	ala	his	gly	leu	phe	320	pro	trp	cys	gly	leu
	TTT	GTT	CAG	ATG	CCG	GCC	CAC	GGC	CTA	TTC		CCC	TGG	TGC	GGC	CTG
	leu	leu	asp	thr	arg	thr	leu	glu	val	gln	330	ser	asp	tyr	ser	ser
	CTG	CTG	GAT	ACC	CGG	ACC	CTG	GAG	GTG	CAG		AGC	GAC	TAC	TCC	AGC
340	tyr	ala	arg	thr	ser	ile	arg	ala	ser	leu	350	thr	phe	asn	arg	gly
	TAT	GCC	CGG	ACC	TCC	ATC	AGA	GCC	AGT	CTC		ACC	TTC	AAC	CGC	GGC
	phe	lys	ala	gly	arg	asn	met	arg	arg	lys	360	leu	phe	gly	val	leu
	TTC	AAG	GCT	GGG	AGG	AAC	ATG	CGT	CGC	AAA		CTC	TTT	GGG	GTC	TTG
370	arg	leu	lys	cys	his	ser	leu	phe	leu	asp	380	leu	gln	val	asn	ser
	CGG	CTG	AAG	TGT	CAC	AGC	CTG	TTT	CTG	GAT		TTG	CAG	GTG	AAC	AGC

FIGURE 47 (cont.)

390
 leu gln thr val cys thr asn ile tyr lys ile leu leu leu gln
 CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG CTG CAG

 400 410
 ala tyr arg phe his ala cys val leu gln leu pro phe his gln
 GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT CAT CAG

 420
 gln val trp lys asn pro his phe ser cys ala ser ser leu thr
 CAA GTT TGG AAG AAC CCA CAT TTT TCC TGC GCG TCA TCT CTG ACA

 430 440
 arg leu pro leu leu leu his pro glu ser gln glu arg arg asp
 CGG CTC CCT CTG CTA CTC CAT CCT GAA AGC CAA GAA CGC AGG GAT

 450
 val ala gly gly gln gly arg arg arg pro ser ala leu arg gly
 GTC GCT GGG GGC CAA GGG CGC CGC CGG CCC TCT GCC CTC CGA GGC

 460 470
 arg ala val ala val pro pro ser ile pro ala gln ala asp ser
 CGT GCA GTG GCT GTG CCA CCA AGC ATT CCT GCT CAA GCT GAC TCG

 480
 thr pro cys his leu arg ala thr pro gly val thr gln asp ser
 ACA CCG TGT CAC CTA CGT GCC ACT CCT GGG GTC ACT CAG GAC AGC

 490 500
 pro asp ala ala glu ser glu ala pro gly asp asp ala asp cys
 CCA GAC GCA GCT GAG TCG GAA GCT CCC GGG GAC GAC GCT GAC TGC

 510
 pro gly gly arg ser gln pro gly thr ala leu arg leu gln asp
 CCT GGA GGC CGC AGC CAA CCC GGC ACT GCC CTC AGA CTT CAA GAC

 520 530
 his pro gly leu met ala thr arg pro gln pro gly arg glu gln
 CAT CCT GGA CTG ATG GCC ACC CGC CCA CAG CCA GGC CGA GAG CAG

 540
 thr pro ala ala leu ser arg arg ala tyr thr ser gln gly gly
 ACA CCA GCA GCC CTG TCA CGC CGG GCT TAT ACG TCC CAG GGA GGG

 550 560
 arg gly gly pro his pro gly leu his arg trp glu ser glu ala
 AGG GGC GGC CCA CAC CCA GGC CTG CAC CGC TGG GAG TCT GAG GCC

 564
 OP
 TGA GTGAGTGTGTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGC

 CTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACACCTGCGTTTTCACTTCCCCAC

[illegible][illegible]

FIGURE 48

Motif -1
 Ep p123 ...LVVSLIRCFFYVTEQQKSYSKT...
 Sp Tez1 ...FIIPILQSFFYITESSDLRNRT...
 Sc Est2 ...LIPKIIQTFYCTEISSTVTIV...
 Hs TCP1 ...YVVELLRSSFFYVTETTFQKNRL...
 consensus FFY TE

Motif 0
 Ep p123 ...KSLGFAPGKLRLLPKKT--TFRPIMTFNKKIV...
 Sp Tez1 ...QKTTLPPAVIRLLPKKN--TFRLITNLRKRFL...
 Sc Est2 ...TLNFNHSMRIIPKKSNEFRIIAIPCRGAD...
 Hs TCP1 ...ARPALLTSRLRFIPKPD--GLRPVNM DYVVG...
 consensus R PK R I

Motif A
 Ep p123 ...PKLFFATMDIEKCYDSVNREKLSTFLK...
 Sp Tez1 ...RKKYFVRIDIKSCYDRIKQDLMFRIVK...
 Sc Est2 ...PELYFMKFDVKSCYDSIPRMECMRILK...
 Hs TCP1 ...PELYFVKVDVTGAYDTIPQDRLTEVIA...//...
 consensus F D YD

Motif B
 Ep p123 ...NGKFYKQTKGIPQGLCVSSILSSFYYA...
 Sp Tez1 ...GNSQYLQKVGIPQGSILSSFCHFYME...
 Sc Est2 ...EDKCYIREDGLFQGSSLSAPIVDLVYD...
 Hs TCP1 ...RATSYVQCQGIPQGSILSTLLCSLCYG...
 consensus G QG S

Motif C
 Ep p123 ...PNVNLLMRLTDDYLLITTQENN...
 Sp Tez1 ...KKGSVLLRVVDDFLFITVNKKD...
 Sc Est2 ...SQDTLILKLADDFLIISTDQQQ...
 Hs TCP1 ...RRDGLLLRLVDDFLLVTPHLTH...
 consensus DD L

Motif D
 Ep p123 ...NVSRENGFKFNMKKL...
 Sp Tez1 ...LNLSLRGFEEKHNFST...
 Sc Est2 ...KKLAMGGFQKYNKA...
 Hs TCP1 ...LRTLVRGVPEYGCVV...
 consensus G